SEQUENCE LISTING

<110> Daniel Cohen
Ilya Chumakov

<120> Treatment of CNS Disorders Using D-Amino Acid Oxidase and D-Aspartate Oxidase Antagonists

<130> 101.US5.REG

<140> 10/051,681

<141> 2002-01-16

<150> 60/261,883

<151> 2001-01-16

<150> 60/305,445

<151> 2001-07-13

<150> 60/345,211

<151> 2001-10-22

<150> 60/333,881

<151> 2001-11-19

<160> 26

<170> Patent.pm

<210> 1
<211> 86592
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<222> 38288..40388
<223> 5'regulatory region

<220>
<221> exon
<222> 40389..40670
<223> exon Z

<220>
<221> exon
<222> 42666..42778
<223> exon A

<220>
<221> exon
<222> 43415..43519
<223> exon B

<220>
<221> exon
<222> 61159..61402
<223> exon 1

<220>
<221> exon
<222> 64050..64711
<223> exon Ulong

<220>
<221> exon
<222> 64478..64711
<223> exon U

<220>
<221> exon
<222> 68126..68261
<223> exon V

<220>
<221> exon
<222> 69488..69690
<223> exon 2

<220>
<221> exon
<222> 71942..72056
<223> exon 3

<220>
<221> exon
<222> 73962..74038
<223> exon 4

<220>
<221> exon
<222> 74701..74766
<223> exon 5

<220>
<221> exon
<222> 77478..77532
<223> exon 6

<220>
<221> exon
<222> 78762..78866
<223> exon 7

<220>
<221> exon
<222> 81507..81589
<223> exon 8

<220>
<221> exon
<222> 83181..83298
<223> exon 9

<220>
<221> exon
<222> 84879..84977
<223> exon 10

<220>
<221> exon
<222> 84980..85423
<223> exon 11

<220>
<221> exon
<222> 84906..85541
<223> exon 11long

<220>
<221> misc_feature
<222> 86543..86592
<223> 3'regulatory region

<220>
<221> allele
<222> 6199
<223> 24-1443-126 : polymorphic base C or T

<220>
<221> allele
<222> 44814
<223> 24-1457-52 : polymorphic base A or C

<220>
<221> allele
<222> 72600
<223> 27-93-181 : polymorphic base A or G

<220>
<221> allele
<222> 84046
<223> 24-1461-256 : polymorphic base A or G

<220>
<221> primer_bind
<222> 6074..6094
<223> 24-1443.pu

<220>
<221> primer_bind
<222> 6559..6579
<223> 24-1443.rp complement

<220>
<221> primer_bind
<222> 44763..44782
<223> 24-1457 pu

<220>
<221> primer_bind
<222> 45257..45277

<220>
<221> primer_bind
<222> 72459..72478
<223> 27-93.rp

<220>
<221> primer_bind
<222> 72763..72780
<223> 27-93.pu complement

<220>
<221> primer_bind
<222> 83791..83811
<223> 24-1461.pu

<220>
<221> primer_bind
<222> 84307..84324
<223> 24-1461.rp complement

<220>
<221> primer_bind
<222> 6180..6198
<223> 24-1443-126.mis

<220>
<221> primer_bind
<222> 6200..6218
<223> 24-1443-126.mis complement

<220>
<221> primer_bind
<222> 44795..44813
<223> 24-1457-52.mis

<220>
<221> primer_bind
<222> 44815..44833
<223> 24-1457-52.mis complement

<220>
<221> primer_bind
<222> 72581..72599
<223> 27-93-181.mis

<220>
<221> primer_bind
<222> 72601..72619
<223> 27-93-181.mis complement

<220>
<221> primer_bind
<222> 84027..84045
<223> 24-1461-255.mis

<220>
<221> primer_bind

<200> 84047..84065
<201> 24-1461-256 mis_complement

<200>
<201> misc_binding
<202> 6187..6211
<203> 24-1443-126.probe

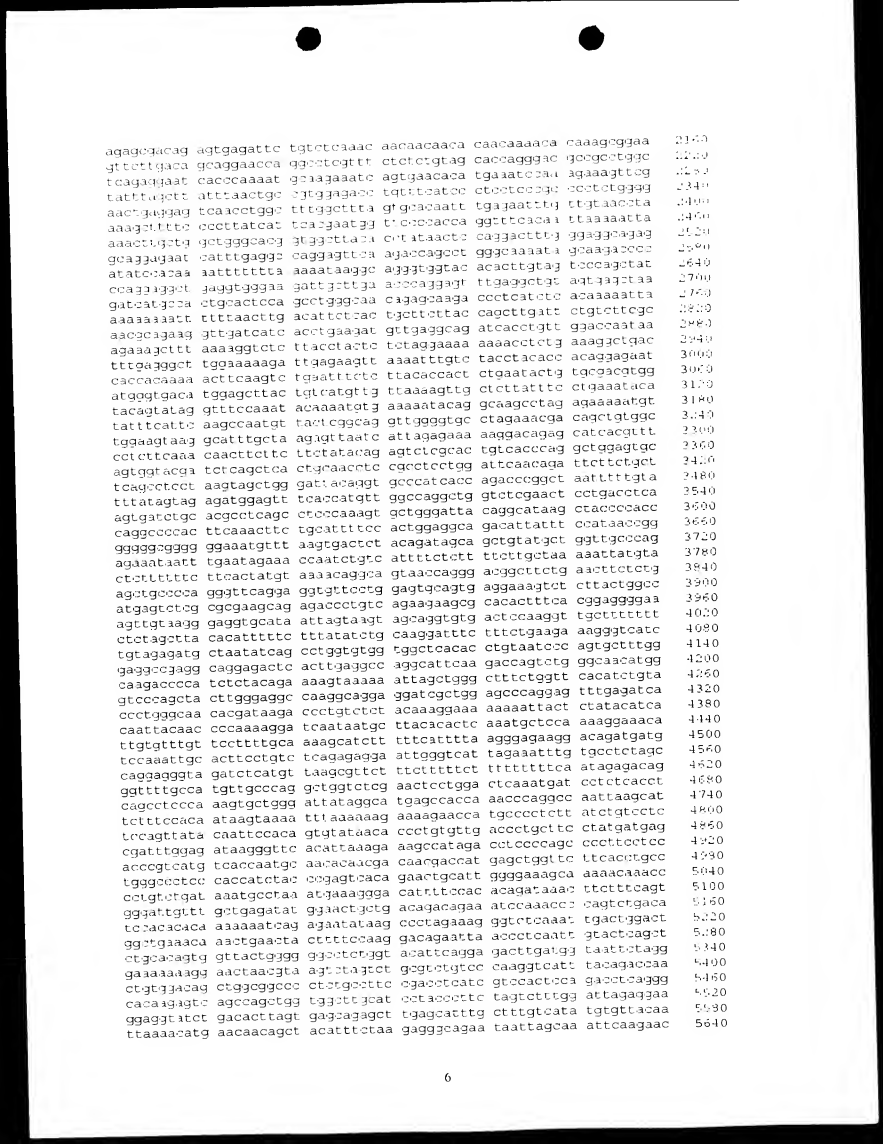
<200>
<201> misc_binding
<202> 44802..44826
<203> 24-1457-52.probe

<200>
<201> misc_binding
<202> 72586..72612
<203> 27-93-181.probe

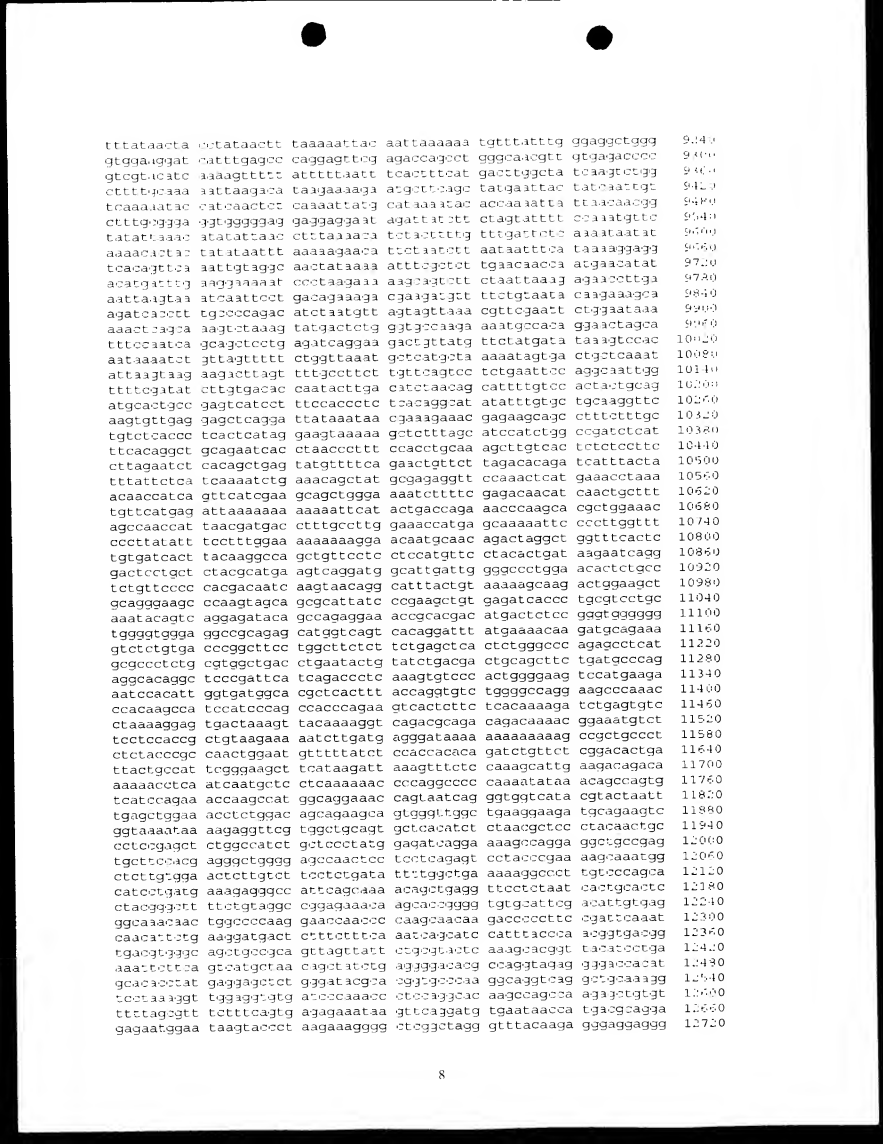
<200>
<201> misc_binding
<202> 84034..84058
<203> 24-1461-256.probe

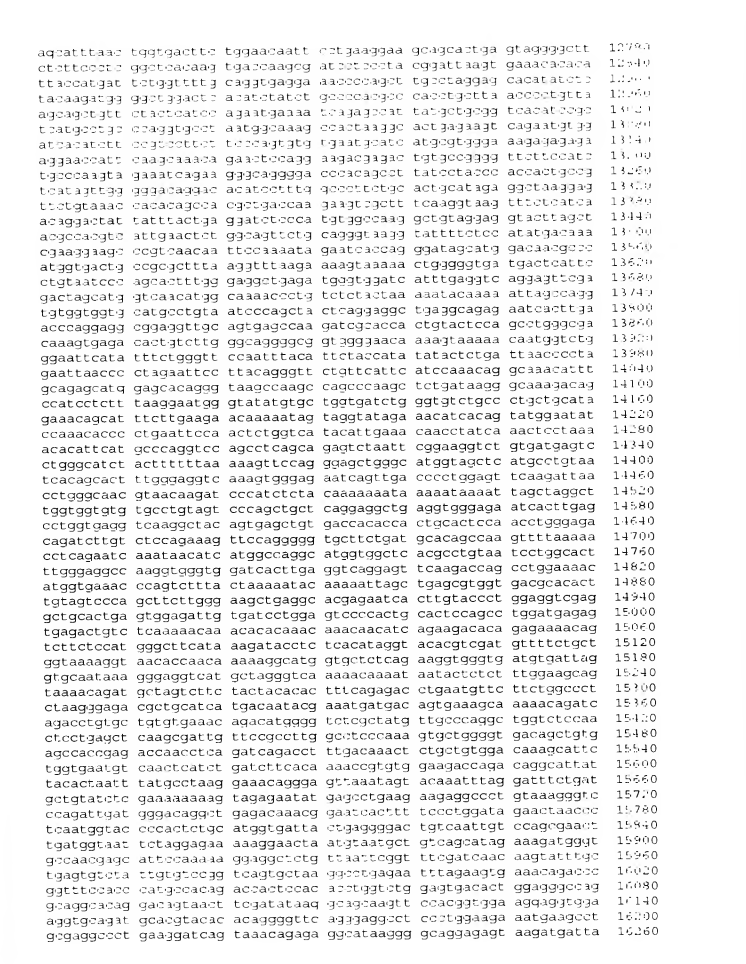
<400> 1

```
attattggaa caggccacac ttgcgaggga agtccctgcc tcagaaagat tcagaaaagc     60
tagacagtca ctggaagaac aattacaacc gcaagacggt caaacactaa acaccgctat    120
gcctcagaac cgtacagata atggccaaat agatgggggt ctgggcattt ctgagagcac    180
ctgactggtg gcacccccatc ctaatggacc atgccctcca gtctccaagt ggctcttcag   240
agctcacatc cgaacacctc ctatgctaca ggttcttcta gccccaggtt cccaaccacc    300
ccaaggccac agaggccagc cccaactcca tcttctacat gtgtcacagg aaacttttct    360
atagtgctat ttattatgta ctgcgggggt gggggccatg tcataaaaga aatgtcctcc    420
cttttttatt catctccttc taacaagcat caaagtctca gtcgctagca tgtgacttac    480
agaagctctc atgggaacaa gacaagacca tactgttacc gtgacactca cggcctccct    540
gactggtttc tgctgttgat tctgcctcaa atgctcctca aatgcacctt gctgctccgc    600
ctccacccta gagctcgcct gactgcccac ttgcccgtta agagtcggct taggcttcac    660
tcctgccaga aaggtcctgc caggtgctct caacagtcac ccctcctgt ggtctcacaa     720
aaccccagca cctctcggtc actctctccc tcctatctgg ttgtgactgt cttccatgct    780
cacttagaag ctctctgagg ccaagaactg tgtgtactgt tgcttctttg tttacctggg    840
cctagcccat tgcctcatac acaggagaat gcaaataaat catatgctta atgaatgagt    900
cgatgaatga atgatgaata aagggaatct aatctagttt taacaaatcc aggttttgca    960
atgatctcac aggcattcat ttatcttgtg atgtcagggg agtgactcca ccctcatttc   1020
acacgcatct tgcggtcaat gctctaactt acttggcctc cagttagtgg gaaattacaa   1080
gctacacttc aagcctctga ctaggacctg ccatgaagta cttgggaatc agtggagtat   1140
cactgtgggg tgaggtgtct gaggcgaggc ccaccaatct ccatacttct ccccgggccc   1200
ttctgcctga gagggtctcc ctgcttccct tggcagactc tggttttgcc ttctgggttc   1260
ggcgttgttg tcacctcctt caggaagcat ttctggctaa ggtgccccac tctatagcag   1320
ctggtgtaaa acctctctaa gcaaacagca taacttttcg tcctctcaat tgactctgag   1380
ttctgagagc acagcctgga gctggcacgg tgcctggcac agagagctga aatggcacac   1440
cctagtgttc ccagtggctc gactccccag gctctccatc aggacgcagc cctctcccac   1500
cttctgatgga tatgggacca tggaatgctt tgtccagcag caactctttgc ctcccctaca  1560
ctaagggaaca cctagcccat cagactcacc tttccttact ggaaaaagtcc actccagca   1620
agatatttctc ctcggtgtcc tggcgcccgc tgctgtacac caccaccatg taccggaccc   1680
ggtccgccca ggcgctctcc aggcgcactg cctggaacag ggcagacatg ctctcactaa    1740
cctgcctttg gaggtggtgc ctcccctcca tctccaatgc aagatcaaca ctttcagtgt    1800
tctacctttc cctctgggag ttaaaaatga agagaaaatt ctttggctggg catggtggtt   1860
caggcctgta atcccagcac tttgggaggc caaggtaggc agatcacttg aggtcaggag    1920
ttcgagacca gcctggccaa catggcaaaa ccccatctct tactaaaaat acaaaaatta    1980
gctgggcatg gtggcgtgcg cctgtaatcc cagctactca ggggactgag gcacgagaat    2040
ctcttgaacc cgggaaggcg aggttgcagt gagctgagat catgccacca catgccagcc    2100
```

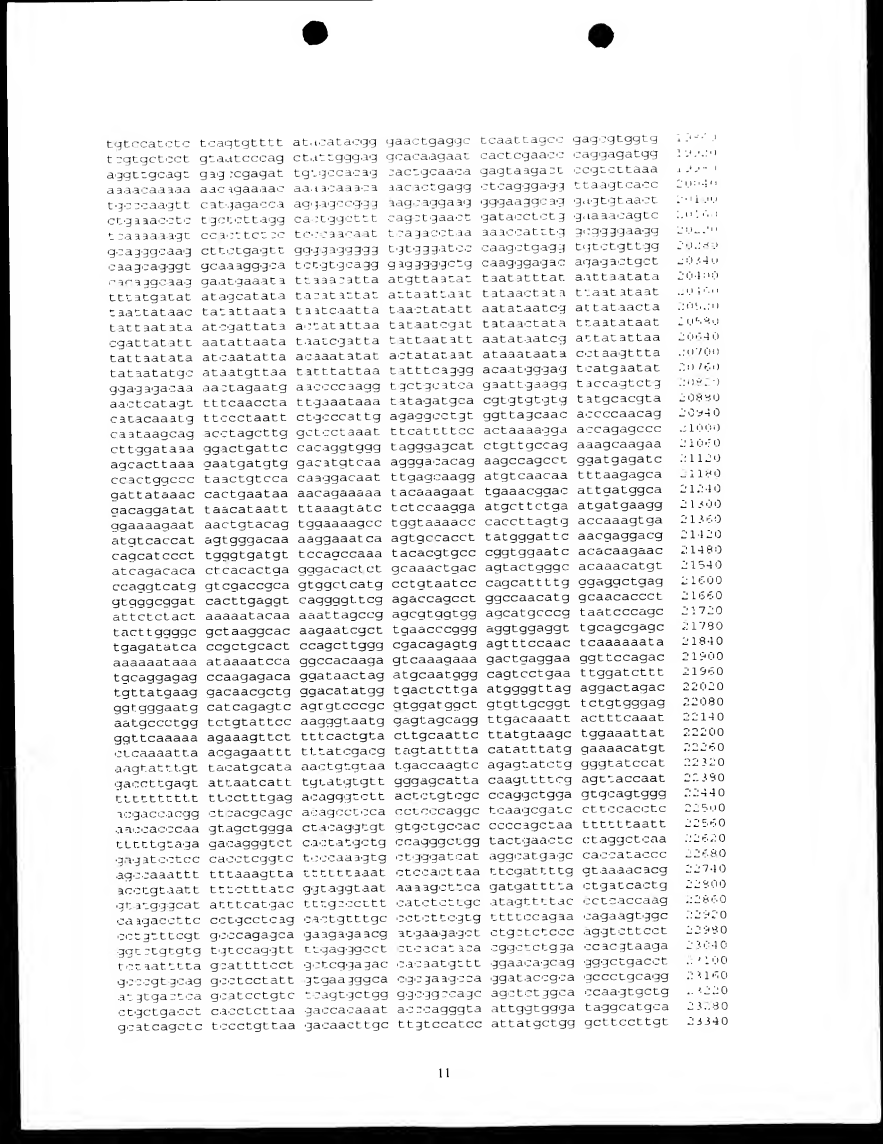


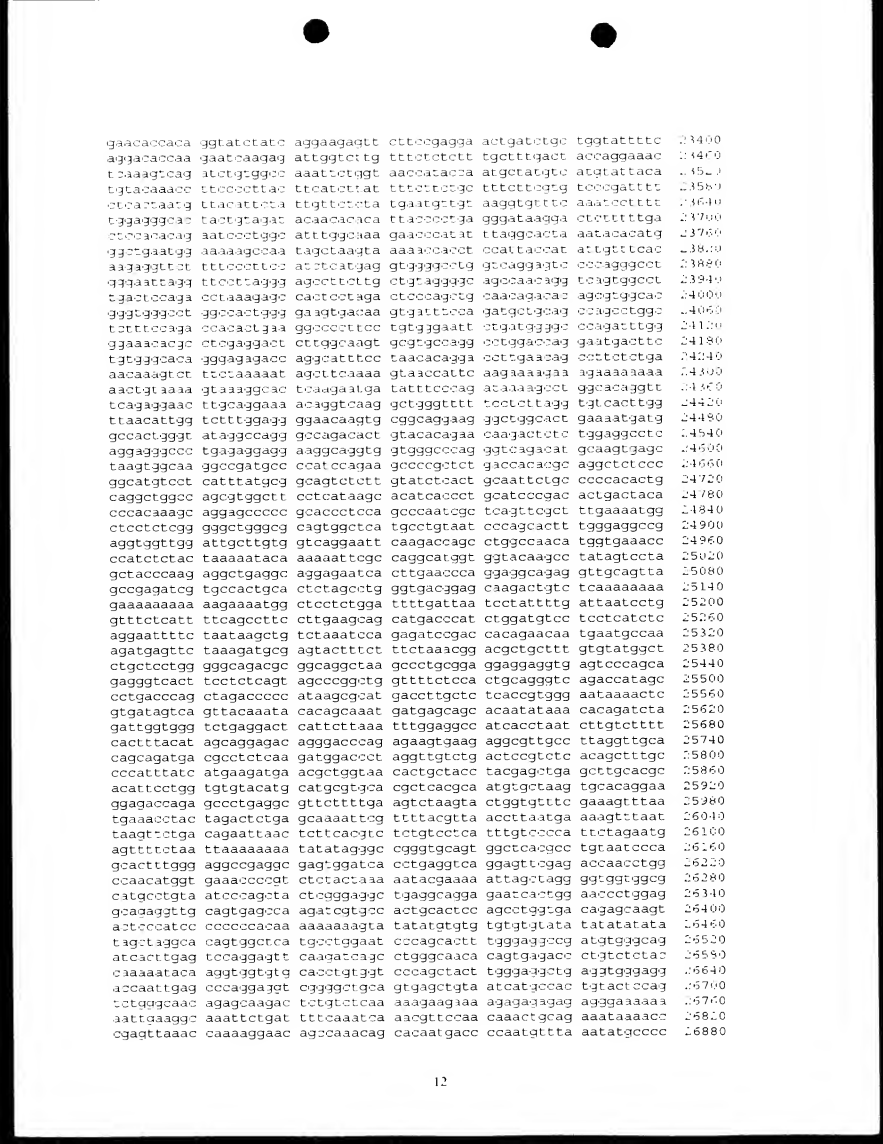
gaagaatctg gctgagtatg gcacctcaaa cctataatcc caatgctttg agaggctgag 5700
gtggggagat ggcttgaggc caggagttgg agaccagcct gagcaacata gtgagtgaga 5760
cctcatctac acacacacac acacacacac acacaactag ctgggtgtgg tgacacgcat 5820
ctatggtccc agctactcag ggaggctaagg ctggacaatc acttgagccc aggaggtcga 5880
ggctgcagtg agctatgatc aggccactgc acgccagcct gggcaagaga atgagatccg 5940
tctcttaaaaa aactttttat ataattaaaa taaaaaaaag aatgaagggt ctgtttatag 6000
ctgtattgta ctagaagtca tcgtaataac aatgatagtt acccatatat atatacagca 6060
cctactacag gtaggtatgt tcacgcgata actcttaaatt tccatattgt ctgaggcacc 6120
agtatttgat gcccattgta atgactagga aactgaggct tagaagtcga cctgttacgg 6180
cttagtaagt tggagaacta ggatcagaag acaggttctg ctggcttcaa aacaaatact 6240
atttccazaa accacactgc ctcctgtgac aggacagtta ttttctttgc ttaaaacaga 6300
cctaaatatt atcaacatca gtatgtgaaa atactgactg agccttggtg tttgctataa 6360
attgcatggt gtagaatttt aacctgagca ctcagatcta aaatgaagct gaatgacttg 6420
aggttaaaca aacaaaatgt ttacaagaaa actggccaca atagctggtt ggtttcacct 6480
gctgctgttc tgaaaggtaa aggccttctc agctcacaga cattcaatta tgcactgcct 6540
ctccaagaaa tgccctgaga tgctgtccac ctacgacaaa gatccactta catgcaagca 6600
cttttcctc tttcttctt tttgagatag ggtcttttc ttttgtcacc caggctggag 6660
tgcagtggcg caatcgtggc tcactgagca acacagtgag caactatagtg agacctcatt 6720
tacatacaca cccacaaaaa actagctggc tgtggtgaca catcagcctc gacctcctgg 6780
gctcaagcaa tcctcccacc tcagcctcca acttgctgg gattacaggc atgcgccacc 6840
acgccagct aattttgta tttttgcag agatagggtt tcactgtgtt gttcgggtta 6900
gtctggaact cctgggttca agcgagatct gcccacttg gcctcccaaa tcctggaatt 6960
acaggcaaga gccaccgtgc ctggccataa gtgtgttttg ttgttattgt ttttaagaaa 7020
cagagtctct ctctgtcacc caggctggag tgcagtggcg tgatcctagc ttgctgcagc 7080
ctcaaactcc tgggctcaag cgatcctcc aactcagcct cccaaagcac tgggattaca 7140
ggtgtgagat accatgcagg gccacgcaag catttcttga attcctcttt ctaactgcct 7200
tcagctctga gtcaagtctc ctaagaaaac cagtcttact acttagtagg cacttcttat 7260
ttaaactcag tttgatcctc acccttattac ttctgtctac ttcctaaaaa caaactatta 7320
cagaatcaag acttcctact acagtgtcta tctcagagtt ggagccaaag gccctcaag 7380
aaatttctcca aatgagtgtt tttcaaatgc ttggagaaat ccatcccaag attaggtata 7440
cagcactcca gatggttatt ttcaagtgga cgacatctgg ctataattca ttttggtgca 7500
tttgttaaaa agtcaggctg taacttacag cctgcaatta actgataaac tacagagagg 7560
aaatctttgc atcccagcag gatgctgctg accttactcc tgacgcagac agacatgaca 7620
taaaaggttg gaaaatgtgc gtggtctgct caagagagag catctgagcc tctgcctgca 7680
ctggtcactg caaacctgcg tccactatgt ctaaggcctt caaactcagc aacatcacca 7740
acaatggaag tttcctctgc tgtccagaaa agaagctcca atgtaagagt atcaacttag 7800
agccctcacc tgcatgcttg tgggggtgct gaagactccg ctggccttga gggctgcttc 7860
cctgttgtaa gaagaggget gcgcctttca ccatgaaaaa gctctcactt aagctggtaa 7920
ggataagacc agagcacagt tagaccggaa ttcagacagg aaaatggaca aagaattact 7980
gcagggaaa aagctttagc gtggacaaat ggcatgtaaa atgcaaatag gatgaaactg 8040
cttttataat aattccacgt agtactttttc tcaaacctg ctttttgctaa aagcttgctg 8100
ctggagaatt ttcgtgacaa aataatgctt ctgtgacaac acccaaagtt ctacataggc 8160
tctccagggc ccctttctgc agaatactgg acagggatct cactgtcata taacatttttc 8220
ttctttcttt ttttttgag acgcagtttc actctgtcat ccaggctgga gtacagtggt 8280
gtgatctcaa ctcactgcaa cctctgcctc ctgggttcaa gcgattctca tgtctcagcc 8340
tcccagaag ctgagattac aggcatgtgc caccatggcc agctaatttat tgtattatta 8400
gtagagacat ggtttcacca tgttggccag gctggttcta aactcctgac ctccagcaat 8460
ctgcctgctt aggcctcctg gagtgctgtg attacaggcg tgcaccacgc cagccataac 8520
attttctaag aaaagagaac aactccctga ttaggagagg gcagtctact ttgtgaattc 8580
tcatgctctt gctgttgatc tctgcttcta actctctggc ttttaacaac tccattgttt 8640
ctttggtgact tccctgatg gaatacaagg atgaaattac actttcacta gttgtttgca 8700
ttttaagaaa agtgggagg ggccgggtgg ctcaagccg tatcccagc attttgggag 8760
gccaaggcag tggatcactt gaggtcagga gttcgagacc agtctggcca acatggtgaa 8820
acctgtctc tactaaaaat gcaaaaatta gccaggcgtg gtggcacatg cctgtaatcc 8880
cagctactca gaaggctgag gcacaagaat cgcttacttg agcccaggg acggaggttg 8940
gagtgagcca agatcgcacc actgcactcc gcactccagc ctgggcgaca gagcaagact 9000
gtgtctcaaa aaaaaaaaa gaaagaaaga aagaaagaaa aaagtgggtg gatactgact 9060
tgtgatttaa cttagtcaag gttgtcctgt ccactatttt tgaggaaaac ctcaagttgg 9120
cccaatgaat ttctcagcag aatgaatctt tggcctttgt tattttagct agcaataaca 9180

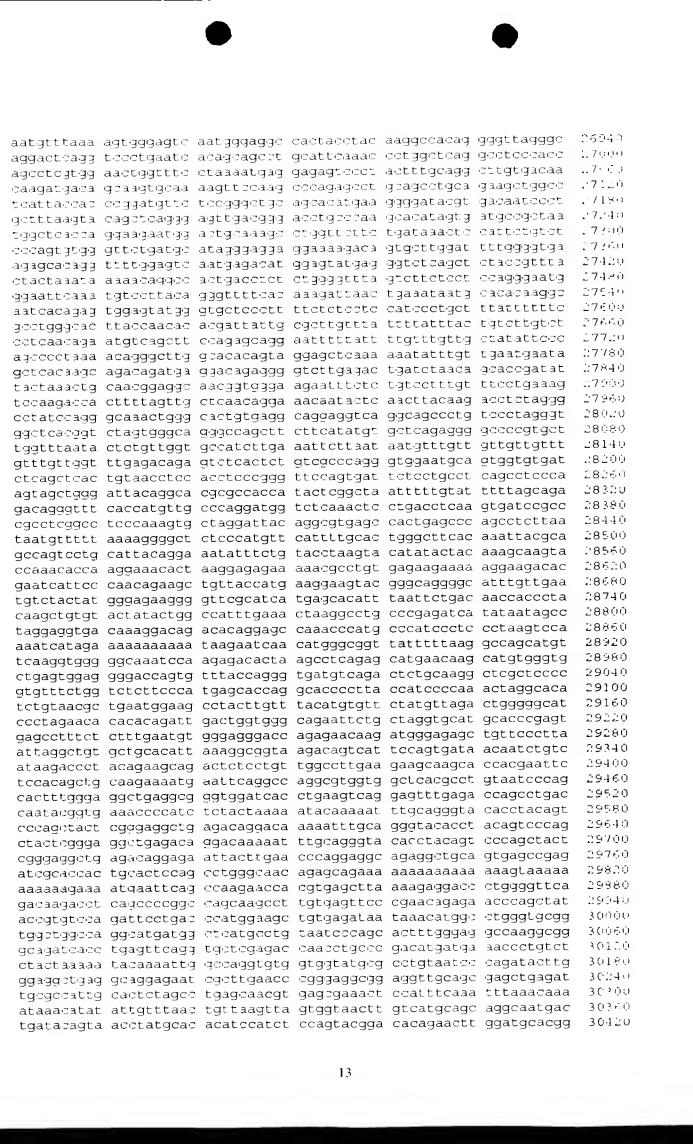


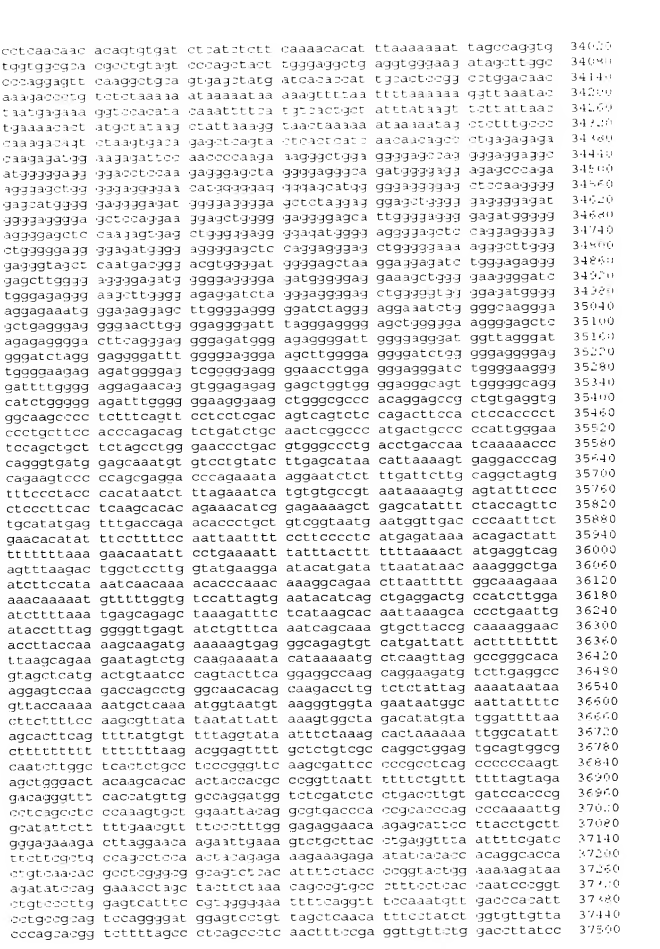


tgctacatgt acttattgt gaacccagga ggatttggcc tctgtcataa aaggccccct 16320
tgtgggttca taaacctcaa tttacaaatt gtgctttata tatcagttcc ttttaagttt 16380
ggttagcgga aattggtttc ttagaacttg atcatccctg agtgaactca caaattcaag 16440
tttcagaatg tgcaaacata agaaacaaac ctcatgcttg tggttgagac atcgcaatgt 16500
caacatcaca aattctcagc acctgaatgc ctggtatact atcaacatat attgttttaa 16560
atatgtaaat aatagctttc tagttataga gagtttgtcc ctacattttt ccacttaatt 16620
tttacaatcc cattcccctg atgaaacaac cccagcctgg gcaacatggc aaaatcctgt 16680
ctctacaaaa aatacaaaaa ttagctgggt gtggtggcgt gcacctgcag tcccgggtat 16740
ttgggaggct gaagtgggag aatcacttga gcccgggagg cagaggttgc agtgagccaa 16800
gattgtacca ctgcactcca gcctgggaca cagagggaga ctctgtccca ccccccccg 16860
cctcccaaaa gaaaagaaaa gaaagaaaag aaaatgaaac cacaaagact gggagaagat 16920
aaatgacttg tctgtggtca tctggctaat aagaggtaga atgggggctga aaaagttcgg 16980
tgctcttcct gaagaatcca taggtcagaa agcagcacca tctgacctgc agcaatagca 17040
gcaacgtgga aagctaatca actgacctca aaaccacttt cagtgaggct ctggatggat 17100
tcagaacccc aggcctagca aagtgaagtt gataaagatg taaaggagat cgaaaattca 17160
ccatttggag agagattagc taaagactgc aggtcggatg gaaaatttct tccatggttc 17220
tcccacaggt tcttccctca tttggaactc gtgtttaaaa gtcacaaaga ccctgagttg 17280
ggccaaggtc tcgttcttct tcactgtggg ccttgcagtg caacatggca gggcctcgtt 17340
ccaaatgtca ctcttcagag cctaagaaaa caagtaactt tagggacaca cctgtcaacc 17400
ggagctccca aattgtaccc ccctaaacac ataatgctga gcatagaaaa attccagctc 17460
tgcagagcgt tatacttagg gaaaggggtc acagacaagg aatgctggca gggctcatta 17520
caaatatctt tgctgctgga acatgtattg tttggctaga agcgtaggc ttctctcaga 17580
gagaaggaat gtccaaaagt atttcagaca gtaagagaca ttctctgagc cagctacaca 17640
gctctcttc aaaccaacgg gtagcggcaa gcagctgaac tgaccagcga gctcgcaaaa 17700
gcaagctttt ttttttttt ctccctaaat aagacagcaa gtgatgtgtc ttggcttggt 17760
ttagcaaatt ttaagatagt tccctgatga ccccaagagc cctcaggccc catggaagct 17820
ggagctaatg catcttcctc caagcatcat ctgctctacc aggatctaag ccccttcacg 17880
agggcagaag gtataaaggc tgcactgtgc gggaaatgct atggcagcaa agacagccaa 17940
acacgccaga aataacaggc acatgaagga aatgtttctg agacagctca aaaatttcga 18000
gaagagatta tcccgactgt cccaggttct cagccctgtc tatggtatgc agccccatac 18060
cacagtcatt tgtcaccgag tcctaacttt gtcagaggcc cctccttcca ggtctctcag 18120
gcaccaccca gttctggccc tcctcacccc cgtgagccag gcgacatcca agcagcccca 18180
cggtgcaccc ggctctgtgc tgcatttctt gaatgtcctt gaaggccagg gctgttgtat 18240
tctctaccca gttctgtctt agtatgggag ccactggcca gatgtgttga ctgaacactt 18300
aagatgcagt aagtgtgacc aagaaactgg gtttgtcatt ttatttcatt ttagttaatt 18360
taaatttaag tttaattagc tacatgaggc tatcagctgt ggtatagagc agcagagctc 18420
cggaagcttt tggcctggtg agaagaatca ggacaagcgc ctccctggcc tctcgcccac 18480
tctgcacagc cgctaacctt tgctctcatg acattctttc ccagccccag aactttagc 18540
catgtgacat catctattga ttagagtcca aacttcttgt gctaactctc tatgggttgc 18600
cacaattagc cattgtatgt cgttaaccta aatttcattc atctgctatg tcctgacctt 18660
aggggcttag aatatagtta gaaaacagta tttcagaata aaaaaccatt cttgtattac 18720
ctctcgcact attccccctg ttctccatgc ttcgcattct ctgttctatc cccagctata 18780
gcactgtccc cataaagcct actgtggttc tcggctcatg gtgtccttcc tcccatctgc 18840
ctcccgacgt catgcctgtc ttccagtgtt tatgccttct ccaggaagct ttctcttgtg 18900
gccctcgctg tgagctatag ctcctccctt tcaacatctt ctagcacctc ctcttactgt 18960
gcaggtgagg acactgaggc ttaagggtta agtcacttgc tcaaggtcac atacagtcgg 19020
tcctctgtat ccgcgggttc ctcatcagtg gattcaacca actacagaca gaaaatacag 19080
tatttgaggg atgctgaact ctttgaatta gtgggttctg cgggtgctta agcatccatg 19140
gattttgtta tcctcggcaa aggcgggggt cctgaaacca atccccttgg atactgaagg 19200
aaagaccacc cttagtgata ggaacctagg aacccaagtt ccctcatttc caaatcgtgt 19260
tccctgaccc actttattac taactagtgg tgaagccatc ttcctgccag tatatttaa 19320
cttcacaatg ggatgtgagg gccaggatgc acatgctttt taaatctccc tctgtggttg 19380
acatacagta gattgaaagt aagtgctggt agatacactg gccaagctgt gctcttctct 19440
gaagtcagta ttccaggagt aactcacccc ggtcatctct gtgcccttgg cacactgggc 19500
actcccaca cacaggttga acctggcaaa taaqactcac agtatcatgc cacgtgcgag 19560
ttaaagccac ctggaggtca ggtcaggtct tcctgacaac tgagtgcttc aaataaacaa 19620
acagtagcta agttccccac atcacttga gtgtctggag agctaggcct atgactttc 19680
tgtctagga tcctctcag tgcccagaaa acagtggaca tcaataaatg taacaccaat 19740
aacatcttcg ttgagcgcta tgctaagcac atcaggtatg ttaactcatt tattccccag 19800



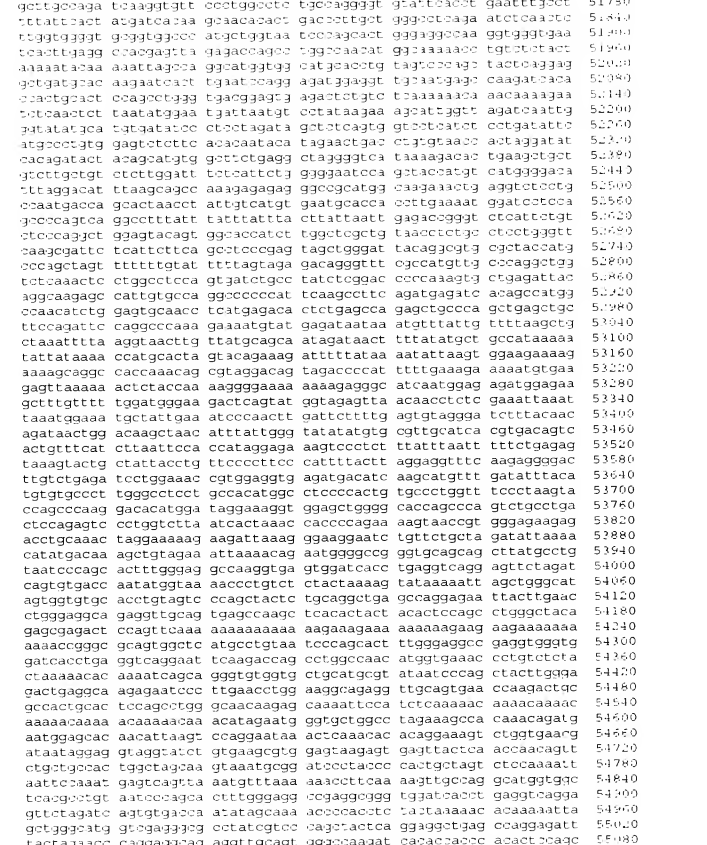



aatgtttaaa agtgggagtc aatgggaggc cactacctac aaggccacag gggttagggc 26640
aggactcagg tcccagaatc acagcagcct gcattcaaac cctggctcag gcctccaccc 26700
agcctcgtgg aactggttcc ctaaaatgag gagagtccct actttgcagg cttgtgacaa 26760
caagatgaca gcaagtgcaa aagttccaag cccagagcct gcagcctgca gaagctggcc 26820
tcattaccac ccgjatgttc tccgggctgc agcactgaaa ggggatacgt gacaatccct 26880
gctttaagta cagctcaggg agttgacggg acctgcccaa gcacatagtg atgccgctaa 26940
tggctcacca ggaagaatgg attgtaaagc ctggttcttc tgataaactc cattctgtct 27000
cccagtgtgg gttctgatgc atagggagga ggaaaagaca gtgcttggat tttggggtga 27060
agagcacagg ttttggagtc aatgagacat ggagtatgag ggtctcagct ctaccgttta 27120
ctactaaata aaaacaqqcc actgacctct ctggggttta gtcttctcct ccagggaatg 27180
ggaattcaaa tgtcctcaca gggttttcac aaagattaac tgaaataatg cacacaaggc 27240
aatcacagag tggagtatgg gtgctccctt ttctctcctc catcctctgt ttatttttc 27300
gcctgggcac ttaccaacac acgattattg cgcttgttta ttttatttac tgtcttgtct 27360
cctcaacaga atgtcagctt ccagagcagg aatttttatt ttgtttgttg ctatattccc 27420
agcccctaaa acagggcttg gtacacagta ggagctcaaa aaatatttgt tgaatgaata 27480
gctcacaagc agacagatga ggacagaggg gtcttgagac tgatctaaca gcaccgatat 27540
tactaaactg caacggaggc aacggtggga agaatttctc tgtcctttgt ttcctgaaag 27600
tccaagacca cttttagttg ctcaacagga aacaatactc aacttacaag acctctaggg 27660
cctatccagg gcaaactggg cactgtgagg caggaggtca ggcagccctg tccctagggt 27720
ggctcacggt ctagtgggca gggccagctt cttcatatgt gctcagaggg gccccgtgct 27780
tggtttaata ctctgttggt gccatcttga aattcttaat aatgtttgtt gttgttgttt 27840
gtttgttggt ttgagacaga gtctcactct gtcgcccagg gtggaatgca gtggtgtgat 27900
ctcagctcac tgtaacctcc acctcccggg ttccagtgat tctcctgcct cagcctccca 27960
agtagctggg attacaggca cgcgccacca tactcggcta atttttgtat ttttagcaga 28020
gacaggtttt caccatgttg cccaggatgg tctcaaactc ctgacctcaa gtgatccgcc 28080
cgcctcggcc tcccaaagtg ctaggattac aggcgtgagc cactgagccc agcctcttaa 28140
taatgttttt aaaaggggct ctcccatgtt cattttgcac tgggcttcac aaattacgca 28200
gccagtcctg cattacagga aatatttctg tacctaagta catatactac aaagcaagta 28260
ccaaacacca aggaaacact aaggagagaa aaacgcctgt gagaagaaaa aggaagacac 28320
gaatcattcc caacagaagc tgttaccatg aaggaagtac gggcaggggc atttgttgaa 28380
tgtctactat ggagaaggg gttcgcatca tgagcacatt taatttctgac aaccaccta 28440
caagctgtgt actatactgg ccatttgaaa ctaaggcctg cccgagatca tataatagcc 28500
taggaggtga caaaggacag acacaggagc caaacccatg cccatccctc cctaagtcca 28560
aaatcataga aaaaaaaaaa taagaatcaa catgggcggt tatttttaag gccagcatgt 28620
tcaaggtggg ggcaaatcca agagacacta agcctcagag catgaacaag catgtgggtg 28680
ctgagtggag gggaccagtg tttaccaggg tgatgtcaga gaagcaagca ccacgaattc 28740
gtgtttctgg tctcttccca tgagcaccag gcacccctta ccatccccaa actaggcaca 28800
tctgtaacgc tgaatggaag cctacttgtt tacatgtgtt ctatgttaga ctgggggcat 28860
ccctagaaca cacacagatt gactggtggg cagaattctg ctaggtgcat gcacccgagt 28920
gagccttttct ctttgaatgt gggagggacc agagaacaag atgggagagc tgttccctta 28980
attaggctgt gctgcacatt aaaggcggta agacagtcat tccagtgata acaatctgtc 29040
ataagacctt acagaagcag actctcctgt tggccttgaa gaagcaagca ccacgaattc 29100
tccacagctg caagaaaatg aattcaggcc aggcgtggtg gctcacgcct gtaatcccag 29160
cactttggga ggctgaggcg ggtggatcac ctgaagtcag gagtttgaga ccagcctgac 29220
caatagggtg aaacccctc tctactaaaa atacaaaaat ttgcagggta cacctacagt 29280
cccagctact cgggaggctg agacaggaca aaaatttgca gggtacacct acagtcccag 29340
ctactcggga ggctgagaca ggacaaaaat ttgcagggta cacctacagt cccagctact 29400
cgggaggctg agacaggaga attacttgaa cccaggaggc agaggctgca gtgagccgag 29460
atcgcaccac tgcactccag cctgggcaac agagcagaaa aaaaaaaaaa aaagtaaaaa 29520
aaaaaagaaa atqaattcag ccaagaacca cgtgagctta aaagaggacc ctgggggttca 29580
gacaagacct cagccccggc cagcaagcct tgtgagttcc cgaacagaga acccagctat 29640
acctgtctca gattcctgac ccatggaagc tgtgagataa taaacatggc ctgggtgcgg 30000
tggctggcca ggcatgatgg ctcatgcctg taatcccagc actttgggag gccaaggcgg 30060
gcagatcacc tgagttcagg tgctcgagac caacctgccc gacatgatga aaccctgtct 30120
ctactaaaaa tacaaaattg gccaggtgtg gtggtatgcg cctgtaatcc cagatacttg 30180
ggaggctgag gcaggagaat cgcttgaacc cgggaggcgg aggttgcagc gagctgagat 30240
tgcgccattg cactctagcc tgagcaacgt gagcgaaact ccatttcaaa tttaaacaaa 30300
ataaacatat aattttaac tgttaagtta gtggtaactt gtcatgcagc aggcaatgac 30360
tgatacagta acctatgcac acatccatct ccagtacgga cacagaactt ggatgcacgg 30420

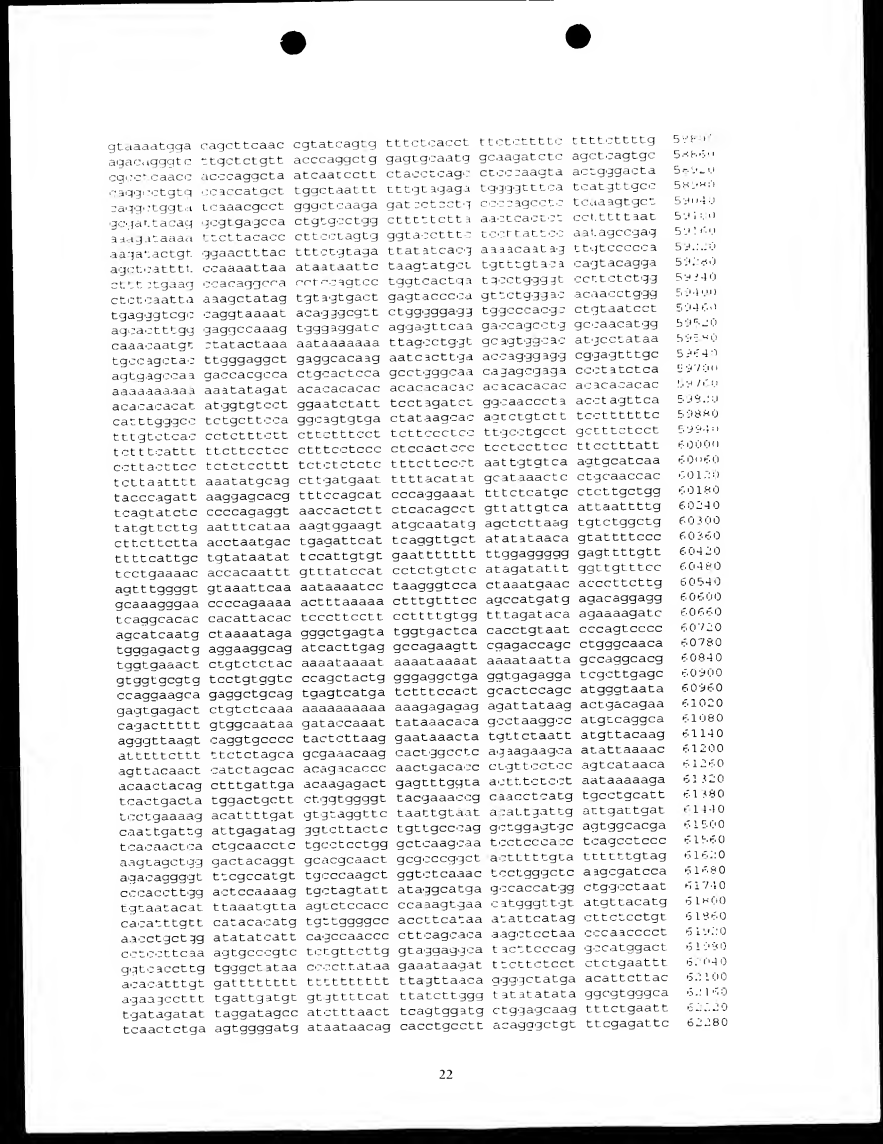


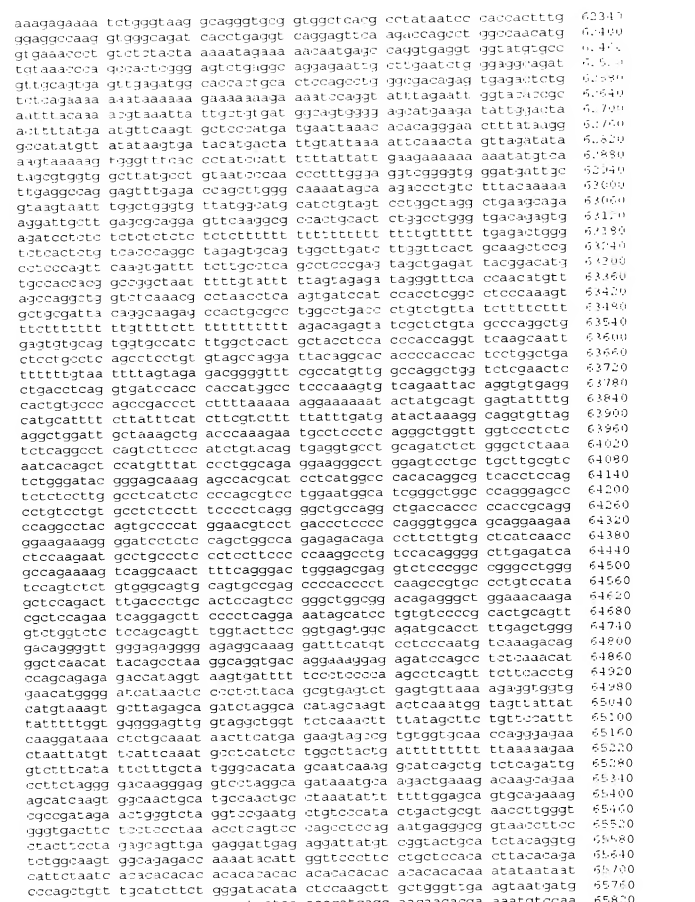
tgttttttctc ttttaagggg aggggtgtat gtttaaagag aatccacttc ctccgcagag 37560
ccaggcaata acagctgagt gatgaacacc attttcaaaa aacgaacccа ggcaagactt 37620
gcacagtgga aggtggccag gaatcaggcc gtctgtttgt gggtcttgaa agctcttgat 37680
ggttctcgaa aagattaaa catttgatac gaaacatcct aggctatcgg tttatttata 37740
taaatgcaag aaagagatat ttaatatttt ctgaaatcta aaaggccacg agtttgggct 37800
ccagaagtac ctatgactta tttttatttt ttttctttca gagagcaaac tgaaaataag 37860
aaggaaacac atacacaccc ccaaaacaac tccgcaccgc tgggacttgg catgttttct 37920
atgttgcaca gaggcgccca ttgaatggga aagagaaacc tggaaagctg tgatggctgg 37980
gagagatgca gggctgatcg aggacagaaa tgaggcagga gccaagggcg aaggaaaaag 38040
ggtccagaga taatgtagga aggggcctgg gcagcaagga acacccacca ggaggtggca 38100
acttcaacca agaatgagta caccagcccg gcgcagtggc tcacacctgg aatcccagca 38160
ctgcaggagg ccgaggtggg cggatcacct gaggtcagga gttcgagacc agcctgccca 38220
acaaggtgac acgctgtctc tactaaaaat acaaaaatta gccaggcacg gtgacatgta 38280
cctgtaatcc cagctacctg ggaggctgag gcaggagaat cacttgaacc caggaggcgg 38340
aggttgcagt gagccgagat tgcaccactg cactccagcc tggtgaaaga gcaatacttc 38400
gtttcaaaaa aaaaaaaaaa gcgtacacca gaggccctag gagtccctac atcatattaa 38460
gatgaagtac atacaagatt ctgcagaggc acctacccca cactgaagga gaagtggaaa 38520
gagcagggaa ggcttcaatg accacaaaac aagtcacaag agcaacaaat tgacaaagag 38580
tatgttgggg tctaacccag tggttctcaa tgcggagcaa gttctctcaa caggagacat 38640
tggaagatgt ctggaaggcat ttttctggag gtcactactg gcacctagtg ggtagagcac 38700
aggatcccac aacacacagg atggtccccc cacaagagag aatgttctgg tcccaagtat 38760
caacagtggg ttgagaaact ctggtccaat ccaaaaaaat gcctggagat ctgcccatga 38820
aaattagtca ctttgaatgt ttctcagaaa taacaatgtt atgatccatt cctgaaaatt 38880
attgatttat ctattcttgt gctctgcctg ttcacacaaa ggaactgaga taagattcac 38940
aggaaaatga cataagtgac ataatcaaag actgggaaaa aaagaaaatt aaaagagaaa 39000
agagagcctt aggactgagg gctgtgatcc cccgtttccc acgccggcag caggcctggc 39060
tgtgtcagga aagcactgcc ctaagtgtct gactcattat gaagttgcaa tttgaagagt 39120
gatgacgtga cttgggggta cggacttcac aatcatttaa ctctcggtca ctctctgagg 39180
ttctcagatg aaaggccatc tcaggtcagt tattccagag aaactacatc tgccaaggaa 39240
cacatgaaag aggtaattca gtccttttag atgagccagg gcccacacac aggaagcaac 39300
tcaagcgagg gcggaccagg gcagaaccgg cctggcctag gtctcctgac cccatacaca 39360
cttgctgtct ccatccccac ttgcttctca cctcaacaca tctgaacgag ggccttggct 39420
tcgggaaaca tcccagcgca ttcaaagcca agcaatgaat gctgcagctt tgctatgatc 39480
aaatataaagc tgcctgagtt ttactttatg tttatcagag ctttaggcac ttggtaaaaa 39540
taatgctttta tataataatt gaaaatgtat ctagtagaca caacacaaat gtccaacaaa 39600
atgtggtaca tccatacaat ggagtattat acagccatga aaaggaatga agtactgcca 39660
catgctacaa tatgcatgaa ctttgaaaac gtgatgctga gtcaaagaag ccagacacaa 39720
aaggccacac agggtgtgat tccatttata taaaatgtcc agaataagca aatccatagg 39780
tacagaaagt agattagtgg ttgcctaagg ttagggagaa tgggggcagg ggaggctgca 39840
ggtgagggct aacgggtaca gggttcctat tttgggggag atgaaaatgt tctggaatta 39900
gatggtggtg gttgcacaat ctcatgtata tactaaaaac cactgaattg tacacttaa 39960
aatggcaaat ttatggtatg taactaaaaa atcaataaac cttaaaatgc gtaagacaag 40020
aacagattag gttgcaagta actctaggaa catggggttt tgaatcagaa atctgggctg 40080
acaggttcag cctgaagcca acctctcccc ttacctcaca taaacttttg tgatgagaac 40140
ctgagattaa gtgcataaat tgccagacag cagtgacgcg aacagaaaac agcccctcct 40200
cgctgtggga aggaaaggcg ctcctgcaac ctaacttctc agtagcaggc tattgatcgc 40260
cagtgttctt ttcctctaa tcagcgtgta gaggggggatt actagaacct tctgtgtata 40320
gataactcat gaatggcctc tcctctccaa ggaggggggt gtgaaggttc aacttcccag 40380
ccactctgaa aatgtccctg ccaatcccag caaaacaagg ctgaagaact accctaccag 40440
gagacagggc tgtcaaggca aatgcaaaca ttattctctt gtctctctca gacacacaaa 40500
cctcccccgt gttatcagtc aacttcccac actcctccc acaaagaaag ggagctgaaga 40560
gcccagatgc tggctgcgga acttcctggg cctgggacag cagggccgct cctccagtct 40620
tctctaaaca cagctaaggg tctgcaggcg gacactcagc cttgttatag gtaagagttt 40680
agaccagagg ccttgacggg ttcttcaaga gatggtgggc aagattgcgc gacctagagg 40740
tcatcctgc agctacatg ggctgacctg ctcagaggac caaggcccca ggctaggaac 40800
agccaggcca acctgcagg ctaagaggg aacagtgccc tcaatcaacc ccagaggaaa 40860
aaqtggccag gcaaacggac ctgggccaca cacajaccca caaaaacgcg cacagtgcca 40920
ggacacgcaa cccaggaatg cacctatgca atcacccaga atgggtcaca gccacaaga 40980
aagatagatg cacataaaca cacaggcctg agtgatgtta cagaaaggaa aagccagact 41040

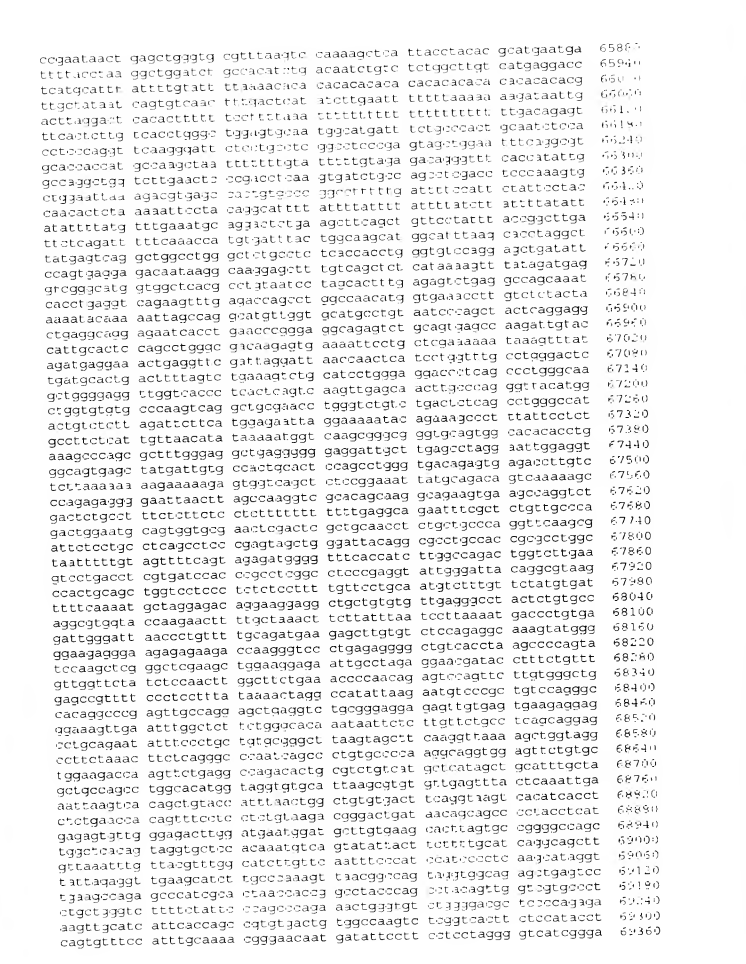
cagggctggag tgcagtggca tcatcatagc tcacagcagc ctcgacctcc agaggctcaag 48180
ggattctccc acctcagcct cctgagtagc tgcgactaca gacacacacc actatgcccg 48240
gttgactttt tttaattatt attatacttt aagttctggg gtacatgtgc agtatgtgca 48300
ggcttgtrac atagatatac acgtgccatg gtggttttgt gcacccatca accggtcatc 48360
tgcattagat attttcccta atgttatccc tcrcctggcc cacaccccc tgactggccc 48420
cgqtgtgtga tgttcccat gcccgttga tttttaaggg ttttgttttt ttqtttgttt 48480
ttttagagac gaggttctca gctgggtgga gcggctcatg cctgtaattc cagcacttta 48540
gaaggtaagg cggggagatt gcttcagccc agggagttcaa gaccagcctg ggcaacatgg 48600
cgaaaccaaa aaatgcaaaa aattaactgg gcatggtggc acatgcctga ggtgaggatg 48660
ggagtatcgt ctgagcctgg gagatcaagg ctgcagtgag ccatatcat gccactgtgc 48720
tccagcctgg ttgatgaggt gagaccctgt gtctaaaaaa taaaagaaat gaaggtcttg 48780
ctgtgtttcc taggctgttc ttgaactcct aggctcaagc aatcctcctg cctcagccac 48840
cccagttgct tggattacag gcacaagcca ccatgtccaa tcctggcaat gtcttgattt 48900
tagacttctg atctctacaa ttgcaagaga ataaatttat gttgttttaa gccacgaaat 48960
ctctggcaat ttgttacagc agccatacga aatgaatata aaactcaacc tccatttggg 49020
ctttaaaaaa catatcatta taatgccatt acccaggtgc ttcccaagcg 49080
ttgtgtcatt ttctcattca ctcaactcat ccaataaact atgtttgttg ctctcctgag 49140
cactagtcta ggaatctggg ttccatcagt gaacaaaatg gaatcactgc ccttgaagag 49200
catccaatca agtggaaaat atagtaaaaa tatatatata tgcaaatatg tttaaaatca 49260
tatgtggtaa atatattgca tttaaatgaa ttaataggcc ggacacggtg gctcatgcct 49320
gtaatcccag cactttggga ggccgaggcc agtggatcac ttgaggccag gagttcgaga 49380
ccagcctggc caacatggcg aaaccccgtc tctactaaaa gtacaaaaat tagccagttg 49440
tggtggtggg tgcctgtaat cccaggtact cgggaggctg aggcacaaaa atcgcttgaa 49500
ctgagggggt gcggaggttg cagtgagccg agatcatgcc actgcactcc agcctgggtg 49560
acagagtgag actgtctcaa aataataata ataataatta attaaatgaa ttaatattgg 49620
taagggtcct tagaacaaga taggcactga tatgtgtcaa ataaatgaaa tatgatgtcc 49680
aatcatgaaa aagcttggga gaaaaacaaa gcaggctaag ggcagagtaa tggagggagc 49740
cacttagaca aatggtcagg gaagcttctg ggtgaggtga tatttgagca gaggaatcac 49800
catgacagca ccaccaggga ggtgtagaaa ccctgggatc tgcctggttc attcaaactg 49860
gcctcccac taaggaactg tgaggtactt tttctgagac ccatttttctt tctgtctgtg 49920
tcacccaggc tggagcgcag tggcgcgatc tcggctcact gcaacctcct ccccccaggc 49980
tcaagtgatc ctcccacctc agcctcctga gtagctagga ttacaggtgt gtgccaccat 50040
acccagctaa tttttgtatt tttagtagag tcggcgtttc accatgttgg ccaggccagg 50100
ctgccacctt ggcttcctac agtgctggga ttacaggtgt gagccttcag acccagccga 50160
gacccactgt ctttctctgt aaaattgata tgaaagtgat agtgctcggc cgggcatagt 50220
ggctcacgcc tgtaatccca gcactttggg aggccaaggt gggcagatca cctgaggtca 50280
ggagttcaag accagcctgt ccaagacggt gaaaccctgt ctctactgaa aatacaaaaa 50340
ttagccaggt gtggtggtgg gtgcctataa tctcagctac tcaggaggct gaggcaggag 50400
aatcgcttga acccaggaag cagaggttac agtgagtcga ggtcccgcca cttcactcca 50460
gcctggacaa caagcaaga ctccatctca aaaaaaaaaa aaaaaagaaa agaaaagaaa 50520
agaaagtggt agtgctgacc tcagagcttg gttgtgtcaa ttgaacagca tactatgcag 50580
gaaaggcaga gcgtgctgtc ctatttacta atagtaccta aggtattgag ttgaattgtg 50640
tcccacaaaa attcactagt ccctgtgaat gggacttat ttggaaatga ggtctttgca 50700
gctgatcaag ttaagatgag gtcattaggg cgggcccta ttcgcatatg actgtgtccg 50760
tatgaaaagg gggaaatttg ctgggcgcgg tagctcatgc ctataatccc agcactttgg 50820
gagaccaagg cgggtggatc acctgaggtc aggagttcga gaccagcctg ccaacacatgg 50880
agaaaccctg tctctattac aaatacaaaa ttagccaggc gtggtggtgc atgtctgtaa 50940
tcccagctac ttcggaggct gaggcaggag aatcacttga acccgggggg tggtggttgc 51000
agtgaactga gattgcgcca ttgcactcca gcctgggcaa caagagcgaa actgcatctc 51060
aaaaaataaa caaacaaaca aataaataaa taaataataa aaggggaaat ttggaccctg 51120
agccaaggag aaaatgcttc ctgaaggttg tagttgtgct gccacaagcc aaagagcacc 51180
cgagatggtc agcaaaccac cagagctagg agtgagaagt gaggagcaga tttgcgtggc 51240
cttctgaaga aaccagcaac tcgatttcag agttccaggc tccagaactg agagagtaaa 51300
tgcctgtggt ttaagcctcc cagtttgtgg cactttgtta cagcagccac aggaaaggaa 51360
cgcatctaac atgatcattt catcagctgc agaaaatgag gctcaaagca aggctagggt 51420
ttgaacccag gccaactaga ccccagacca catacatggg ttgttggcct ttctagcctg 51480
ggaggtgaca gtttggatgt tccactattt gcaggaacg gtgttcagag gattcaaagc 51540
ttctgcacc tgggccaggg tgtccagtgt tagatatgga agtcagtgtc ctgggcttc 51600
tgacaaagca tctctctggg tgggtcaata ggcaccagca gccaggcagt tgaggatct 51660

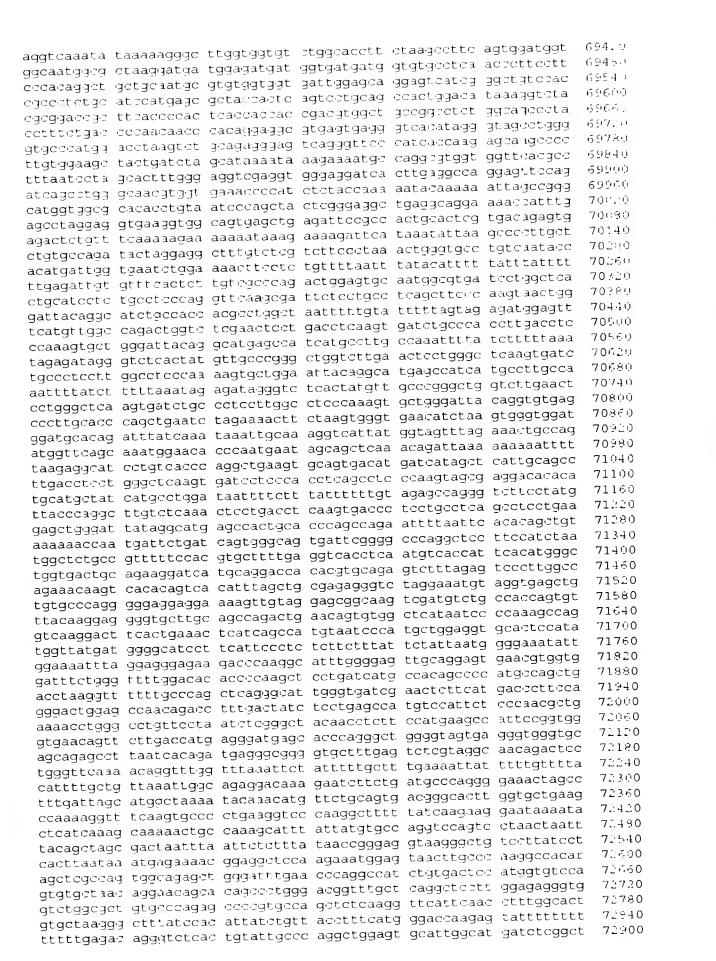


ctaaacatga ctatagccat ataaaagtat gttatаaagc tgggcgtggt ggctcacgcc 55520
tgtaatccca gcactttggg aggctgaggc gggtggatca cctgaggtca ggagtttgag 55580
accagcctga ccaacatgga gaaaccccgt ctctactaaa aatacaaaaa ctagctgggt 55640
atggtggcgc atgtctgtaa tcccagctac tcaggaggct gaggcaggag aatttgcttg 55700
agtcgggagg tggaggttgc agtgagccga gatcgcacca ttgcactcca gcctgggcaa 55760
cagagcgaga actccgactc aaaaaaaaaa aatgttatta aactacacac cactatаaag 55820
aaatgataat gcaaaaatca taaaggacga aaaaaaaaag gaaatajatt taactacaaa 55880
aaagtttttg ttttgctttt gttttttag agttagagtc ttgttctttt tcccaggctg 55940
gtacaatcat agctcactgc cacctttgaac tcttgggctc aagcaatcct cctgccttct 56000
aaaatgcgtt tgcaaaaatt atаactgaga aaacaatgac agtgaaagag atctgaccta 56060
actgactcca tcttgcttct aacctccaag ctgtccgtgt tcattcctgg gtgtaggcca 56120
aactaatttt gggaggaatt tagtttatag tttaacttta tcaaagttta actaagatgt 56180
taatagccca tttttcaaaa caaacccctt tcctgcctgg ggactagact gcctttgcag 56240
gactaacaaa tttatatagc tatcagatta gaaattatgg tttaggagtc atgcagctga 56300
agcctacaag attctgaatc tcccaaattg ctcctggaga taacatcacc attgtaaaac 56360
ctaagatcag tgcttgacat attttgcaga cctcgcactc gatgaatcag ctggcactac 56420
ccaaatggat aaacaggctc atctgatctg tggtccccac ccaggaaatg acccagcata 56480
agaggaccgc ttcaactcct ataacttttgt ctccaacctg aacaatcaac actcccctac 56540
tttctgaccc cctacccacc aaattaccct taaaaacctt agccaggcgc agtggctcat 56600
gcctgtaatc ccagcacttt gggaggctaa ggcaggcgaa tcacctgagg tcagggttcg 56660
aqaccaacca tggccaacat agtgaaaccc catctctact aaaaatacaa aattagccag 56720
gtgtggtagt gtgcgcctgt aatcccagct actcaggagg ctgaggcagg agaatcgcct 56780
gaacccggga cacagaggtg gcagtgagcc aagatcactc cactgcactc cagcctgtat 56840
gacaagagca aaactcggtc tcaaaaacaa aacaaaacaa aaaacccaca gaaaaaaacc 56900
ctgaaccatg atcctaaact ctttcactat tgcagttccc ctgacttgat acattggctc 56960
tgtctaggca gcgggcaagg ataacccatt gggcagttgc acctcagcct cctgagtagc 57020
tgggattaca aatgcaagcc acagctaaaa aaattttcaa acttttgtag gacagacaaa 57080
ttgggaaaaa catttgcaac aaaggggatga tacacataca tataaagagt tcatтcaagg 57140
ctgggcacag tggctcacgc ctgtaatccc agcactttgg gaggccgagg caggcagatc 57200
acgaggtcag gagttcaaga ccagcctggc caatatggtg aaaccccatc tgtactaaaa 57260
atacaaaaat tagccgggtg tggtggcatg cgcctgtaat cccagttact caggaggctg 57320
aggcaggaga attgcttgaa cccgggaagc agagattgca gtgagccgag atcgcaccac 57380
tgcactccag cctgggtgac agagtgagac tccatctcaa aaaaaaaaaa aaaaaaaaag 57440
agttatttca aattaataag aaaaataacc aacacaattc aatagaaaaa tggggaaaaa 57500
agaataggca ctttacaaag aaatаaatac aaagcccagt ggacatgaaa ctttgccatc 57560
tccttagcag ggcttggagt aagatgggga gaaggaagga tgcaaaattt aaggaggctg 57620
tcactctcag ggccatgtaa gtacaaagtg ggcatatgag ggtaagtgcc tccttaaatg 57680
tgtaaattgc tagagccctg ctggatggca gtctggcaaa atggatcaat attttaaatg 57740
tacaaaccct ggcacaatga ttccatttтt aggaactgac cttatggaaa cgatcaggca 57800
agtgtgccaa gaaacacatc taggatgttt ttaatgtcga caaattagaa atgacaggta 57860
aattcaaccc tacggactga ctttaaaaat tgttacatct ggctgggcat ggtggctcac 57920
gcctgtaatc ccagcatttt gggagccaaa catgggagga tcgcttgagc ccaggagttc 57980
aagaccagtc tgggcaacat agggagaccc cgtcgctaca aaaaaaaaaa aaagtaaaaa 58040
ttagccaggt tggtggtgca tgcctgtagt tctaactact caggaggctg aggaggagag 58100
atcacttgag ccctagaggt caagactaca gtgaactgtg attgcgccac tgtactccag 58160
cctgggcaat agagtgaggc cgtgtctcaa aaaagaaaaa aaaatgttac atccaggtac 58220
attggcatcc tgtgtaaaaa ggatgccatc ctgtagtccc agctgcttgg gaggctgagg 58280
caggagaatc gtttgagccc aggaattcga ggcttcagtg agctatgttc acaccactgc 58340
actccagcct aggcaacaga gcaagacttt gtcaataaat taaaagaaaa ataaaaagta 58400
aacagaggaa ccaacatatg tcacatgata ctatttttgc cacctgccg tgtttatgtt 58460
cacatttgga aatatttgcc caaagtaatg gtcectattt cctggtggtg ggattaattg 58520
caggggattc ttacttttct ctttatgcct gctgcataaa tacttgaaat ccttaaatac 58580
tgcttaatac ttgaaaaagt gattaaagct aattttgtct gagtaagaga gtgggagtta 58640
acctgttatt ctgtaacttc ctggccccac cagggttgac tcctgcagag cattctccag 58700
gtaaatgttt ttgcctggc ctgtctgtat ttcagaacta ccaggaggtc gttttgttta 58760
tcaaccacce agtgggggtca aaagaccct taacttctac aattccagcc aaataaacag 58820
aagttgcttt cgaaagtcta gggcctccca ttactaggat cagtgagttt aggacttcag 58880
ggtagtggaa agggccttgg tcccacagag ctgtctcagg gcacttaaat ttcctaagt 58940









aggtcaaata taaaaagggc ttggtggtgt ctggcacctt ctaagcettc agtggatggt 694.0
ggcaatgacg ctaaggatga tggagatgat ggtgatgatg gtggtgcttc acacttcctt 6948.
cccacaggct gctgcaatgc gtgtggtggt gattggagca ggagtcatcg ggctgtctac 6954.
cgccttctgc atccatgagc gctacgactc agtcctgcag ccactggaca taaaggtcta 69600
cgcggaccgc ttaccccac tcaccaccac cgacggtggc gcggtcttct ggtacgccta 6966.
ccttctgtac ccgacaacc cacaggaggc gtgagtgagg gtcactatagg gtaggctggg 6972.
gtgcccatgg acctaagtct gcaggtggag tcaggggttc catcacaag agcagccccc 69780
ttgtggaagc tactgatcta gcataaaata aaqaaaatgc caggcgtggt ggttcacgcc 69840
tttaatccta gcacttggg aggtcgaggt ggqaggatca cttgaggcca ggagttccag 69900
atcaggctgg gcaacgtggt gaaacccccat ctctaccaaa aatacaaaaa attagccggg 6996.
catggtggcg cacacctgta atcccagcta ctcgggaggc tgaggcagga aaatcatttg 70020
agcctaggag gtgaaggtgg cagtgagctg agattccgcc actgcactcg tgacagagtg 70080
agactctgtt tcaaaaagaa aaaaataaag aaaagattca taaatattaa gcccctttgct 70140
ctgtgccaga tactaggagg ctttgtctcg tcttcctaa actggtgcc tgtctatacc 70200
acatgattgg tgaatctgga aaacttcctc tgttttaatt tatacatttt tatttatttt 70260
ttgagatggt gtttcactct tgtcgcccag actggagtgc aatggcgtga tcctgctcta 70320
ctgcatcctc tgcctcccag gttcaagcga ttctcctgcc tcagcttccc aagtaactgg 70380
gattacaggc atctgccacc acgcctggct aatttttgta tttttagtag agatggggtt 70440
tcatgttggc cagactggtc tgaaactcct gacctcaagt gatctgccca ccttgacctc 70500
ccaaagtgct gggattacag gcatgagcca tcatgcctg ccaaatttta tcttttaaa 70560
tagagatagg gtctcactat gttgcccggg ctggtcttga actcctgggc tcaagtgatc 70620
tgcctccttt ggcctcccaa aagtgctgga attacaggca tgagccatca tgccttgcca 70680
aattttatct ttttaaatag agatagggtc tcactatgtt gcccgggctg gtcttgaact 70740
cctgggctca agtgatctgc cctccttggc ctcccaaagt gctgggatta caggtgtgag 70800
cccttgcacc cagctgaatc tagaaaactt ctaagtgggt gaacatctaa gtgggtggat 70860
ggatgcacag atttatcaaa taaattgcaa aggtcattat ggtagtttag aaactgccag 70920
atggttcagc aaatggaaca cccaatgaat agcagctcaa acagattaaa aaaaaatttt 70980
taagaggcat cctgtcaccc aggctgaagt gcagtgacat gatcatagct cattgcagcc 71040
ttgacctcct gggctcaagt gatcctccca cctcagcctc ccaagtagcg aggacacaca 71100
ttgcatgctat catgcctgga taatttttct tatttttttgt agagccaggg tcttcctatg 71160
ttacccaggc ttgtctcaaa ctcctgacct caagtgaccc tcctgcctca gcctcctgaa 71220
gagctgggat tataggcatg agccactgca ccagccaga attttaattc acacagctgt 71280
aaaaaaccaa tgattctgat cagtgggcag tgattcgggg cccaggctcc ttccatctaa 71340
tggctctgcc gttttccac gtgcttttga ggtcacctca atgtcaccat tcacatgggc 71400
tggtgactgc agaaggatca tgcaggacca cacgtgcaga gtctttagag tcccttggcc 71460
agaaacaagt cacacagtca catttagctg cgagagggtc tagaaatgt aggtgagctg 71520
tgtgcccagg gggaggagga aaagttgtag gagcggcaag tcgatgtctg ccaccagtgt 71580
ttacaaggag gggtgcttgc agccagactg aacagtgtgg ctcataatcc ccaaagccag 71640
gtcaaggact tcactgaaac tcatcagcca tgtaatccca tgctggaggt gcactccata 71700
tggttatgat ggggcatcct tcattcccta tcttctttat tctattaatg gggaaatatt 71760
ggaaaattta ggaggagaa gacccaaggc atttggggag ttgcaggagt gaacgtggtg 71820
gatttctggg ttttgacac acccaagct cctgatcatg ccacagccc atgccagctg 71880
acctaaggtt ttttgcccag ctcagggcat tgggtgatcg aactcttcat gacccttcca 71940
gggactggag ccaacagacc tttgactatc tcctgagcca tgtccattct cccaacgctg 72000
aaaacctggg cctgttccta atctcgggct acaacctctt ccatgaagcc attccggtgg 72060
gtgaacagtt cttgaccatg agggatgagc acccaggggt ggggtagtga gggtgggtgc 72120
agcagagcct taatcacaga tgagggcggg gtgccttgag tctcgtaggc aacagactcc 72180
tgggttcaaa acaggtttgg tttaaatttct atttttgctt tgaaaattat ttttgtttta 72240
cattttgctg ttaaatggc agaggacaaa gaatctttctg atgcccaggg gaaactagcc 72300
tttgattagc atgcctaaaa tacaaacatg ttctgcagtg acgggcactt ggtgctgaag 72360
ccaaaaggtt tcaagtgccc ctgaaggtcc caaggctttt tatcaagaag gaataaaata 72420
ctcatcaaag caaaaactgc caaagcattt attatgtgcc aggtccagtc ctaactaatt 72480
tacagctagc gactaattta atttctttta taaccgggag gtaaggcctg tccttatcct 72540
cacttaataa atgagaaaac ggaggctcca agaaatggag taacttgccc aaggccacar 72600
agctcgccag tggcagagct gggatttgaa cccaggccat ctgtgactcc atggtgtcca 72660
gtgtgctaag aggaacagca cagcctggg acggtttgct caggctcttt ggagagggtg 72720
gtctggcgct gtgccagag cccgatgcca gctctcaagg ttcattcaac ctttggcact 72780
gtgctaagag ctttatccac attatctgtt acctttcatg ggacaagag tatttttttt 72840
tttttgagac aggttctcac tgtattgccc aggctggagt gcattggcat gatctcggct 72900

tgtttggaaa tcattgtact aagacatggc tatttctcct aacctgggga cacatgaccc 76500
ttgtccagtc ttttccagga aaaacatgcc ctcaagatgt ttttctatct tgaggaaatg 76560
atggaaatga gatagttcca agggtatgct tcaccttctt tttgacttat ttcctgttct 76620
ttgatgtttt ctagtgtatt tctttctttc ttctttttct ttttttcttt tttgagacag 76680
agtctttgtc tgcacccag gctggagtgc agtggcgcaa tctcggctca ctgcaagctc 76740
cgcctcctgg gttcatgcct ttctcctgcc tcagcctccc gagttgctgg gattacaggc 76800
gcctgccacc acgtctggct aatttttttt ttgtatttgt agtagagacg gggtttcacc 76860
gtgttagccg ggatggtctc aatctcctga cctcgtgatc cgcccgcctc ggcctcccaa 76920
agtactggga ttacaggcgt gagccaccgc gcctggctgt ttctagtgta tttctaatag 76980
tgatagatgt ttttcctatg ggatgtttaa aaggagggtg gatgtcctca gcctacctcc 77040
ctcctcatgc ccggctttcg acaaagggga atttggcact ggtacaactc tcccttctcc 77100
tactctgaat ctcattgcct ttgctgttac aaagcaatgt ggtggtcata ggaagtgctg 77160
ggggctaaga ggcctgggtt tgagttccaa ctcctctcatt gactcactct atggccttca 77220
gcaaggcctt tcccccactc catctgccca acaagggggt tggaccatct ctggtttctc 77280
aaaggagatt ttgtggacca ccagtccagt aggtgctcat gagctgattt gatgacacag 77340
ccatcttctc aagcagcatc ctgtgcaact aacgtccgca gaaggttgtt tgggggaagg 77400
tcctgtgcct acccttcttg gtgggatggg ggcagatagc tgaatactgg gcttttttgt 77460
gtgtttgatc atcccaggtt aactgagagg ggagtgaagt tcttccagcg gaaagtggag 77520
tcttttgagg agttgagttg cagggctgat gcggtggatg gggcagggaa gaagtaggaa 77580
ggcctctgct tcttgctgct gagtcgggag ctccctctcc aggctcctag ggtcccacca 77640
ggcctgcctc agcacccttg ccccagaagc actcaggtat tctgaaggga ggaagtctct 77700
gccttcatgt tggtagtggg aacaaaggaa cactggtatc atggtggcca ttaggagctg 77760
atttatatct gagactcaat gagttttggg tctagagagc tggccgcatt ttctcagtgt 77820
cagctgcact ccaaggtcag aacttggttg cttcctagcc ctaccgacat ctgtgttggt 77880
ctttctgcaa agtccagccc ctcagctgac tcacctctaa agaagcacca ccacccaataa 77940
taatgacagg aaaagccacc atctccaggc accagcaaaa agagctttac tgtatggctt 78000
cattcaatcc cagcatctaa aaccctgctt ggcacaagga aggcgctccg tacatgtagc 78060
tactagtgct atgtcatgaa gactaacctg ctctggtcag gccctgatgg acaccgaaga 78120
tacatggtcg acccaatgca gtcctcattc tcagtcattc actcaggaac aatagtagcg 78180
tcttgcaatg tgtgtgtccc ttaacttact cgtggtgaga gtcactgggg ctgggttggg 78240
gagcttaggg gctcacgatg cgtgcttgag atgagatcat ctcatctgta gacagagctg 78300
gggttccaac gtgtcttctg caaatgtctt ggcagagtag aaggcaagag aataaagtta 78360
aaaggagtca gaaggagaaa gagaactctc tctgcttcct ttctgacttc ttttgggagg 78420
ttccaggaag atttcccca tccaaagaac tgttttacaa ccacttttat attcagagtt 78480
gtgcaggagc ctcataacag cctatgaaca gccatgggca gcctcatttt acaggggcag 78540
ctgagaatta aggaggtaac cagacatttt caaggtcaca cgtcagataa atggcagtat 78600
gaaaatttga agccaggccc ctctgattcc tcattgagac ctctccccac tgttcatcag 78660
ggagtagaca gattgaggt agaagaaagg ggaagagaga acaggggata ccagggtctt 78720
ccccaccttt catccccccac taccctgttg gttgctacca ggtggcaaga gaaggcgcag 78780
acgtgattgt caactgcact ggggtatggg ctggggcgct acaacgagac ccccgctgc 78840
agccaggccg gggcagatc atgaaggtga gtgtgagggt gagacccccta ccttttgtta 78900
ataggaagat cattctgcat gcttatttca tcctcaaga tcatggacaa atcaggaaca 78960
tctgttagag gaaccccccg gactgcaggg aattgacatg taaaaaaaac aaacctgtcc 79020
caccccatt gctctctttc aggatttcct cttgatcgtg aagcatgcat gtatgcgctt 79080
gtacctatgt gtgagcagca tatgcctgta ttgcaataaa aatagcaaac attagagtgt 79140
ttaccaagcg cgagatacag tcctaagcac tttattgtgt ttattattat tattaattat 79200
taattgtgtt attatatta tcattgttat tattattttt gagacaggct atcactccat 79260
tgcccaggtt agagtgcagt atcttgatca tggctcactg tagcctgac ctcccaggct 79320
ccaacttag cctactgagt agctgagact acaggcgcat gccaccacca tgctcagcca 79380
atttttttat ttttgtaga gaaaggattt caccatattg ctcaggctgg tctcaaactc 79440
ctgggctcaa gtgatcccc caccttggcc tgtcaaagtg ctgggattac aggcgtgagc 79500
cactcgctc agcctattgt gttaattaat ttagtgatgg ccacagccct tcgagctggg 79560
tactaccata tcgttattgt catcttacag atgaagaaat tgaggcacag aggagttaag 79620
taacaggcac aagttcacac ggtagtacgc agtgccaatg ggattggaat ccaggcaacc 79680
tggcttaaga gcctgtgcgt gcaagcattg ttcaatgcct cctcttgctg tgtttgtgca 79740
tatgagggta tgtgtgtgtg catgtatgtg tgtgtgtgta tgtaaggata tgtgtgcata 79800
tgtgtgtgtg catgtgtgag ggtgtgtgtg catgtgtgag ggtgtgtgtg catgtgtatg 79860
agggtgtgtg catatgtgat ggtgtgtgca catatgtgag ggtgtgtgtg catgtgtgtg 79920
agggtgtgtg catatgtgtg atggtgtgtg tgcacgtatg tgggggtgat tgtgcatgta 79980

tgtgaggggtg tatgtgcata tgtgtgatgg tgtgcgtgca tgcacaccat gtgaggggta 80040
tgtgtgtgtg catgtgtgtg aaggtgtgtg cgcatgtgag agtgtatgta cgtgtgtgat 80100
ggtgtgtgtg tgtgaggtga tgtatgcatg tgtgtgaggg tatgtgtgtg catgtgtgtg 80160
agggtgtgtg catgtatgta agggtatgtg tacatgcatg tgtgtgaggg gtatatgtgt 80220
ggatgcatgt gagggtgtgt gtgtgcatgt gtgtgagggt ggtgtgtgta gggtgtgtgt 80280
gtgcatgtgc acctgtgagt gttcataggt gtgcaggtgt gtgtgcttct gtgtgtaggg 80340
gtgcgtgtgt gtgttcctaa tgtgggctga tgggtgtaac aatcaaatga gtgactgaag 80400
ataagtctc aaatcatcga ggtttatgga gccagcttga gggcgcatcc aggaaaaacg 80460
cgagtcacag atgcacctgt gactcctttt tccaaagagg ttctcaggag atttagtctt 80520
tatacatttt ctttaaaaaa aaaaaagtga gagaagggtg tagcacgcag agaatgattg 80580
catacttgtg aaactttagt tagtgcccag taaatctaca tttcacataa gatgaaggtt 80640
tgggccaggc gtggtgactc acacctgtaa tcccagcact ttgggaggct gaggcaggtg 80700
gatcacgagg tcaggagttc gagaccagcc tggccaacgt ggtgaaaccc catctctact 80760
aaaaatacaa aaaattagct gggtgtggtg gcgggtgcct gtaatcccag ctactcggga 80820
ggctgaggca ggagaatcgc ttgaacccgg gaggcagagg ctgcagtgag cgagactac 80880
accactgcac tccagcctgg caacagagcg agactgtctc aaaaaaaaaa aaaaaaaaaa 80940
aaattgaagg tttgaaggaa aaaggaatgg aggaagttct gtatctggga agataagctt 81000
gtcattgatg ttatcagtgt ggagtctgtt gaaagggctg gttttgttt acccctagg 81060
gaagaaagcc taactttggt caggtcattg aggggaggga tataatgaga cgtgtcggac 81120
ctcccttccc cccgcagctg tgaactcagc tccaaggttt ctctgggact cctgggccca 81180
agaggggtgc tgttcagtcg gttggggact tagaatttta tttttatttc tcatgtgtat 81240
gcatttacat gtgtgtactg gtgctttttt tcggacatgt gggtgaggag aaacaatgct 81300
tcaggagaca ggtcgtgagt ccaattaggg cagctctttcc tgcagaggc aagcagtcag 81360
gtgcagactt gggccatagt gtcatgagag gtcttataag gaatcagcct ggccactctt 81420
gtcaggacat ctggccacag agggcagcaa ggccagccac attgactcac ctccgctgat 81480
gagactttcc tgccctgaat caacaggtgg acgccccttg gatgaagcac ttcattctca 81540
cccatgaccc agagagaggc atctacaatt ccccgtacat catcccaggg taaaattgga 81600
ctgttctcgg gcagaagagt ggtccccttc atgccctctt catgaccctg ctgcctcccc 81660
caagctcctt actccctgca gttgttccct ttcaatgttt ttatgtactt agctatttttt 81720
tattattatt ttttgagaca gagtttcact cttattgccc aggctggagt gtaatggtgc 81780
gatcttggct cactgcaacc tctgcctccc aggttcaagc aattatcctg cctcagcctc 81840
ccaagtagct gagattacag gtgcccacca ccacatccag ctaatttttt gtatttttag 81900
tagagacagg gtttcaccat gttggccagg ctggtcttga actcctgact caggtgatcc 81960
acctaccctt gcctccaaa gtgctgggat tacaggcgtg agccaccgtg cctggccccc 82020
tttcaatgtt tttagtgagt ttgagctact gaatccctgg gaaggcagac tcagcctcga 82080
ctgaggtcta ccgtgaacat tcttttggat gacaatagtg gtgatgctgg agacaaaggc 82140
agtggatgta atgtggtgac actaaaagtg gtatgtaggt ggctcacgcc tgtaatccca 82200
gcacttgcg agggcaatgt gggaggattt cttgagccca ggagttcaag accagcttgg 82260
gcaacatggc aagaccccgt ctctacaaaa atacaaaaat tagccgggcg tgatggtgta 82320
tgcctatggt cccagctatt cgagaggctg agatggggag attgcttgaa cctgggaggt 82380
tgaggatgca gtgagccatg ttcacaccac tgtactccag cttgggccac agagcgagac 82440
ccatctcaa aaaaaaaaaa aagtggtgtg aatggcaata atggggagtgg gaatgggaat 82500
ggtgattggg gctgatggtg atgataatgt taacggtgga gatgacaatg tcactgaaac 82560
cagtggtggt gttcatggga tgacaatatt gttgatagcg gaatggtggt attagggata 82620
atattgtatt gatggggaag acagcgttca tgggggtggt gattagcgta agagttgtag 82680
agtggtgatg ttaatggagg tggtctggtg ctgatgagga gatcaatgtt gatgaaggtg 82740
tgattgggag tggggatggt agctggtgct gatggaaatg acactatcaa tgatgttaat 82800
actgtagcag agctgacagt ctcaaaggca atgttaataa catggttgca ccaaccatgt 82860
tatctcaatg gcgatgttac tggtgtcgtg gagatgacaa tatcaatggc aatgttagtg 82920
gtggtggtga aatgatgaat gcagttggtg gtgatgacct attaatgata gtagcaaaga 82980
caatgttgtt gatggagatg acaacattga tggaagtggt gatggaagag ttcgttgttg 83040
gtgttgatgg tgatgacagt ggcaattgag gtagtgatgg tggtggtgtt agcagaggtg 83100
acaaggttga tggtaatgac ctttattcat ctcagagcct tcatttttct tcatcttgca 83160
cctcctcaat ttgtatctag gacccagaca gttactcttg gaggcatctt ccagttggga 83220
aactggagtg aactaaacaa tatccaggac cacaacacca tttgggaagg ctgctgcaga 83280
ctggagccca cactgaaggt aaggtaggga ggagtagcag tgccctaaac cgaggtcgtg 83340
ggagcttggt aatgaggaca ctttaggacg ggaagatgcc accgctggga taactgggca 83400
aattaattcc agcaaggat gtggaacata acagaatttg ataatgtaca gggaagttct 83460
tgctatgggc taatgaatcc tgtctggcca tggctgagag cccttggttt tcacatttgt 83520

ctgcgagtga tgatgacagt agtgatggtg atgaggatga gttggtactg atggtgagga 83[illegible]
aaatgctgag aatggtaata gtgatggtga taaggtggtg acagttgtta aaattatggt 83640
ggtggctgat ggtgagggta gtggttgatg atggaattgg tggaaaggtg gaagcagtaa 83700
tggtaatgat gttggtagct gataaagatg gtgttggtgg tagtggtgat tgataaagat 83760
gactgtgatt atattagtgg tggtggtgat gagattctaa aagctaactc cctactacct 83820
aaaaatgtca gcaggaaaaa aaaatccaga aatgagtgat cagcactttt cttttccagaa 83880
tgcaatgatt attggtgaac gaactggctt ccggccagta cgcccccaga ttcggctaga 83940
aagagtactg cttcgcactg gacctccaaa cacaagggta tgctcccatg gcaaggaaag 84000
taatgcccta ttccactcct cagatggctc tggcattttc agggatcagt catgtctgat 84060
ctcaagtttc acacaggctc catagcaggc aaggtcaatg gtggctaata tcccctcctc 84120
tataaatggg gaaactgagg ctcaatgatg gttaaggacc tgctcaaggt tacatagagg 84180
ggcagtggtg atgttaatgg aggtggtgct gatgagatca atgttgataa tggtgtgact 84240
gggaatgtgg atggtagctg gtgctgatgg aaatgacact atcaagtatg ttagtaccac 84300
agcagaggtg acgatctcaa aggcagtgtt aacatggctg cattaactgt ctcattggca 84360
atattaatag tgtggcagag atgacagtat caatggcagt gttaatgatg gtggtgaaat 84420
ggtgaataag gttggttttc taaagtctgt ggtcaaataa caggaaaatg tgtacttact 84480
ggatatgtac ttcgtgtcag acacagcagc aagtccatta catgaatgac cttattaaat 84540
ctcctctgga gctctttggg atagggacag ttctccctat gcttcggatg aggaaactgg 84600
ggtgaattaa gaggtgaagt cacttgccca agtcagacca gtggtggaag gcaggggctg 84660
gatgtgattt gaatttgact ccaaggctat ttccagatat ccatttttgtg gctgccccat 84720
catctcttgc aactgttcca gggggtcccc accattccac cccggtgcca agagaagctc 84780
aggtggcatc tggctttgcc caggactctt cgggaggctc ctgagtcttc cagggcagaa 84840
gaggtttcatc tattcttttc actgtccctc tcggacctgg ccactttctc tcttgcctct 84900
cctaggtcat ccacaactat ggccatggag gctacggact caccatccac tggggatgtg 84960
ccctgtaggc agccaagctc tttgggagaa tcctggaaga aaagaaattg tccagaatgc 85020
caccatccca cctctgaaga ctccagtgac tgctgcctcc ccccacaaga actcccttct 85080
cccctcagcc aatgaatcaa tgtgctcctt cataagccat tgcttctccc tcacttcttt 85140
cctcaaagaa gcatgaggtg agagaaagcc acaaagtcag tgcctggaga agggttcagc 85200
ccaacatggg gccccctcta tcactgaaat ccctctacct tctctgggtc tggcattata 85260
aagaacagct gaggctgtca ttccatgagt cttcagaaga aaggacagct cagaaaatca 85320
aagaggccaa ctgcccagag ccacagaaaa tggaggataa ttgaggctaa gtaacctgat 85380
tacaagttgt actaacatat taaaggttct gaaaagtcct gcagcaaaga caactatctg 85440
atgttgttta acccagtgct tgctaaacct atctggctat ggaactcttt tgcccagagc 85500
acccatgaat gccatgacac aaatctgaga aaatgctgga acagattttg ttgtatctgt 85560
tgtgtttgtt gtaggaggtt atacatacaa ctgggggtgtg gaggggcag agaggtgagg 85620
cactgaacta gtaacacatg gtgtttgttc cacatctaga attccaaatg gcatcagcta 85680
ttcaccgagt ggccccatga gcaccacgta acctttgagg aggggccact ggagggatca 85740
tcccacaagg aaccccttca tagagaactg ttttagtcca ttttctgttg cttatcaacag 85800
aatatctgaa actggagatt ttttttttttt ttttttgaga caggatctca ctctgtcacc 85860
caggctggtg tgcagtggca tgattttggc tcactgcaac ctccgcctcc caggctcaaa 85920
tgatcctccc tcctcagcca cccgagtagc tgggactaca ggcgcttgct accatgccca 85980
gctaattttg tgtgtgtgtg tgtgtgtgtg tgtgttttgt agagagtgtt ttgtagagac 86040
tgggtttggc catgttgtcc aggctggcgt tgaactcctg ggatcaagtg atcctcctgc 86100
ctcagcctcc aaagtggtgg gattataggc ataagccacc acgcctggcg gaaactgtgg 86160
aattaataga gaaaaggaat ttatttatta ccgttataga gtctgagaag tccaaggttg 86220
aggggccaca tctggtgaga gccttctctc tggctggtgc agaggtgggg actctctgca 86280
gagtcccagg gaggcttagg gcatcacgtg gtgagggggc tgattgtgct aatgtgctag 86340
ctcagctctg tccttgtct tagaaagcca ccatttttct tcccaagatg acccattaat 86400
ccatttaact aataacccat taattgataa atggattaat ccatttatga gagcagcgct 86460
cttaggaccc aatcacctct taaaggcgcc acctctccag accaccacta aggtggtgga 86520
ctaaggatta agtctcaacg tgagttttgg cagggacgtt taagcaatag caagaactaa 86580
actcacaaag ca 86602

<210> 3
<211> 1573
<212> DNA
<213> Homo sapiens

<220>

<220>
<221> 5'UTR
<222> 1..143

<220>
<221> CDS
<222> 144..1187

<220>
<221> 3'UTR
<222> 1188..1573

<220>
<221> polyA_signal
<222> 1549..1554

<400> 2

```
tgcactccag tccgggctgg cggacagagg gctggaaaca agacgctcca gaatcaggag      60
cttcccctca ggaaatagca tcctgtgtcc ccgcactgca gttgtctggt ctctccagca     120
gtttggtact tccggctgct gca atg cgt gtg gtg gtg att gga gca gga gtc      173
                         Met Arg Val Val Val Ile Gly Ala Gly Val
                         1               5                  10
atc ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc      221
Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
                15                  20                  25
ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc      269
Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
            30                  35                  40
acc acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac      317
Thr Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp
        45                  50                  55
ccc aac aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat      365
Pro Asn Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr
    60                  65                  70
ctc ctg agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc      413
Leu Leu Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe
75                  80                  85                  90
cta atc tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc      461
Leu Ile Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser
                95                 100                 105
tgg aag gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg      509
Trp Lys Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu
            110                 115                 120
gat atg ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att      557
Asp Met Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile
        125                 130                 135
ctg gag gga aag aac tat cta cag tgg ctg act gaa agg tta act gag      605
Leu Glu Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu
    140                 145                 150
agg gga gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg      653
Arg Gly Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val
155                 160                 165                 170
gca aga gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct      701
Ala Arg Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala
                175                 180                 185
ggg gcg cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc      749
Gly Ala Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile
            190                 195                 200
atg aag gtg gac gcc cct tgg atg aag cac ttc att ctc acc cat gac      797
Met Lys Val Asp Ala Pro Trp Met Lys His Phe Ile Leu Thr His Asp
```

```
        265                 270                 275
cca gag aga ggc atc tac aat tcc ccg tac atc atc cca ggg acc cag      845
Pro Glu Arg Gly Ile Tyr Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln
    280                 285                 290
acc gtt act ctt gga ggc atc ttc cag ttg gga aac tgg agt gaa cta      893
Thr Val Thr Leu Gly Gly Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu
295                 300                 305                 310
aac aat atc cag gac cac aac acc att tgg gaa ggc tgc tgc aga ctg      941
Asn Asn Ile Gln Asp His Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu
                315                 320                 325
gag ccc aca ctg aag aat gca aga att att ggt gaa cga act ggc ttc      989
Glu Pro Thr Leu Lys Asn Ala Arg Ile Ile Gly Glu Arg Thr Gly Phe
            330                 335                 340
cgg cca gta cgc ccc cag att cgg cta gaa aga gaa cag ctt cgc act     1037
Arg Pro Val Arg Pro Gln Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr
        345                 350                 355
gga cct tca aac aca gag gtc atc cac aac tat ggc cat gga ggc tac     1085
Gly Pro Ser Asn Thr Glu Val Ile His Asn Tyr Gly His Gly Gly Tyr
    360                 365                 370
ggg ctc acc atc cac tgg gga tgt gcc ctg gag gca gcc aag ctc ttt     1133
Gly Leu Thr Ile His Trp Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe
375                 380                 385                 390
ggg aga atc ctg gaa gaa aag aaa ttg tcc aga atg cca cca tcc cac     1181
Gly Arg Ile Leu Glu Glu Lys Lys Leu Ser Arg Met Pro Pro Ser His
                395                 400                 405
ctc tgaagactcc agtgactgct gcctcccccc acaagaactc ccttctcccc          1234
Leu
tcagccaatg aatcaatgtg ctccttcata agccattgct tctccctcac ttctttcctc   1294
aaagaagcat gaggtgagag aaagccacaa agtcagtgcc tggagaaggg ttcagcccaa   1354
catggggccc ctctcatcac tgaaatccct ctaccttctc tgggtctggc attataaaga   1414
acagctgagg ctgtcatttc atgagtcttc agaagaaagg acagctcaga aaatcaaaga   1474
ggccaactgc ccagagccac agaaaatgga ggataattga ggctaagtaa cctgattaca   1534
agttgtacta acatattaaa ggttctgaaa agtcctgca                          1573

<210> 3
<211> 1691
<212> DNA
<213> Homo sapiens

<220>
<221> 5'UTR
<222> 1..143

<220>
<221> CDS
<222> 144..1187

<220>
<221> 3'UTR
<222> 1188..1691

<400> 3
tgaactccag tccgggctgg cggacagagg gctggaaaca agacgctcca gaatcaggag     60
cttcccctca ggaaatagca tcctgtgtcc ccgcactgca gttgtctggt ctctccagca    120
gtttggtact tccggctgct gca atg cgt gtg gtg gtg att gga gca gga gtc     173
                        Met Arg Val Val Val Ile Gly Ala Gly Val
                        1               5                  10
atc ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc      221
Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
```

```
                15                  20                  25
ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc      269
Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
            30                  35                  40
acc acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac      317
Thr Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp
        45                  50                  55
ccc aac aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat      365
Pro Asn Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr
    60                  65                  70
ctc ctg agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc      413
Leu Leu Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe
75                  80                  85                  90
cta atc tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc      461
Leu Ile Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser
                95                 100                 105
tgg aag gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg      509
Trp Lys Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu
            110                 115                 120
gat atg ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att      557
Asp Met Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile
        125                 130                 135
ctg gag gga aag aac tat cta cag tgg ctg act gaa agg tta act gag      605
Leu Glu Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu
    140                 145                 150
agg gga gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg      653
Arg Gly Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val
155                 160                 165                 170
gca aga gaa ggc gca gac gtg att gtc aac tgc act ggg cta tgg gct      701
Ala Arg Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala
                175                 180                 185
ggg gcg cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc      749
Gly Ala Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile
            190                 195                 200
atg aag gtg gac gcc cct tgg atg aag cac ttc att ctc acc cat gac      797
Met Lys Val Asp Ala Pro Trp Met Lys His Phe Ile Leu Thr His Asp
        205                 210                 215
cca gag aga ggc atc tac aat tcc ccg tac atc atc cca ggg acc cag      845
Pro Glu Arg Gly Ile Tyr Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln
    220                 225                 230
aca gtt act ctt gga ggc atc ttc cag ttg gga aac tgg agt gaa cta      893
Thr Val Thr Leu Gly Gly Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu
235                 240                 245                 250
aac aat atc cag gac cac aac acc att tgg gaa ggc tgc tgc aga ctg      941
Asn Asn Ile Gln Asp His Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu
                255                 260                 265
gag ccc aca ctg aag aat gca aga att att ggt gaa cga act ggc ttc      989
Glu Pro Thr Leu Lys Asn Ala Arg Ile Ile Gly Glu Arg Thr Gly Phe
            270                 275                 280
cgg cca gta cgc ccc cag att cgg cta gaa aga gaa cag ctt cgc act     1037
Arg Pro Val Arg Pro Gln Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr
        285                 290                 295
gga cct tca aac aca gag gtc atc cac aac tat ggc cat gga ggc tac     1085
Gly Pro Ser Asn Thr Glu Val Ile His Asn Tyr Gly His Gly Gly Tyr
    300                 305                 310
ggg ctc acc atc cac tgg gga tgt gcc ctg gag gca gcc aag ctc ttt     1133
Gly Leu Thr Ile His Trp Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe
315                 320                 325                 330
ggg aga atc ctg gaa gaa aag aaa ttg tcc aga atg cca cca tcc cac     1181
```

```
Gly Arg Ile Leu Glu Glu Lys Lys Leu Ser Arg Met Pro Pro Ser His
                335                 340                 345
ctc tgaagactcc agtgactgct gcctcccccc acaagaactc ccttctcccc                1234
Leu
tcagccaatg aatcaatgtg ctccttcata agccattgct tctcctctac ttcttttctc         1294
aaagaagcat gaggtgagag aaagccacaa agtcagtgcc tggagaaggg ttcagcccaa         1354
catgggaccc ctttcatcac tgaaatccct ctacctttct tgggtctggc attataaaga         1414
acagctgagg ctgtcattcc atgagtcttc agaagaaagg acagctccag aaatcaaaga         1474
ggccaactgc ccagagccac agaaaatgga ggataattga ggctaagtaa cctgattaca         1534
agtttgtcta acatatttaa ggttctgaaa agtcctgcag caaagacaac tatctgatgt         1594
tgtttaaccc agtgcttgct aaacctatct ggctatggaa ctcttttgcc cagagcaccc         1654
atgaatgcca tgacacaaat ctgagaaaat gctggaa                                  1691

<210> 4
<211> 2620
<212> DNA
<213> Homo sapiens

<220>
<221> 5'UTR
<222> 1..1155

<220>
<221> CDS
<222> 1156..1818

<220>
<221> 3'UTR
<222> 1819..2620

<400> 4
gaaacccacg cagcctcctg gattcttccc cgtccctccc tctgtcctgg ggctgtgacc          60
tcctccatgt tattcacagg gtctcagcac gattcatctc aaagcagcga aacaagcact         120
ggcctcagaa gaagcaatat taaaacagtt acaactcatc tagcgcacag acacccaact         180
gacaccctgt tcctccagtc ataacaacaa ctacagcttt gattgaacaa gagactgagt         240
ttggtaactt tctcctaata aaaagatcac tgactatgga ctgcttctgg tggggttacg         300
aaaccgcaac ctcatgtgcc tgcatttcct gaaaagacat tttgatgtag gaagggcctg         360
gagtcctgct gcttgcgtct ctgggatacg ggagcaaaga gccacgcatc ctcatggccc         420
acacaggcgt cacctccagt ctctccttgg cctcatctcc ccagcgtcct ggaatggcat         480
cgggctggcc cagggagccc ctgtcctgtg cctctccttt cccctcaggg gctgccaggc         540
tgaccacccc cacccgcagc caggcctaca gtgccccatg gaacgtcctg accctccccc         600
agggtggcag caggtaagaag gaagaaaggg gatcctctcc agctggccag agagacagac         660
cttcttgtgc tcatcaaccc tccaagaatg cctgcctcc ctcctccccc caaggcctgt         720
ccacaggggc ttgagatcag ccagaaaagt caggcaactt ttcagggact gggagcgagg         780
tctcccggcc gggcctgggt ccagtctctg tgggcagtgc agtgccgagc cccacccctc         840
aagccgtgcc ctgtccatag ctccagactt tgaccctgca ctccagtccg ggctggcgga         900
cagagggctg gaaacaagac gctccagaat caggagcttc ccctcaggaa atagcatcct         960
gtgtcccccc actgcagttg tctggtctct ccagcagttt ggtacttccg atgaagagct        1020
tgtgtctcca gaggcaaagt atgggggaag agggaagaga gaagaccaag ggtccctgag        1080
aggggctgtc ccctaagccc cagtatccaa gctcgggctc gaagctggaa ggagaattgc        1140
ctagaggctg ctgca atg cgt gtg gtg gtg att gga gca gga gtc atc ggg           1191
                Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly
                1               5                   10
ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc ctg cag           1239
Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln
        15                  20                  25
cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc acc acc           1287
Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr
    30                  35                  40
```

```
acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac ccc aac      1335
Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn
45                  50                  55                  60
aac cca cag gag gcg gac tgg agc caa cag acc ttt gac tat ctc ctg      1383
Asn Pro Gln Glu Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu
                65                  70                  75
agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc cta atc      1431
Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile
            80                  85                  90
tcg ggc tac aac ctc ttc cat gaa gcc att cag gac cct tcc tgg aag      1479
Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys
        95                  100                 105
gac aca gtt ctg gga ttt cgg aag ctg acc tcc aga gag ctg gat atg      1527
Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met
    110                 115                 120
ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att ctg gag      1575
Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu
125                 130                 135                 140
gga aag aac tat cta cag tgg ctg act gaa agg tta act gag agg gga      1623
Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly
                145                 150                 155
gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg gca aga      1671
Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg
            160                 165                 170
gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct ggg gcg      1719
Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala
        175                 180                 185
cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc atg aag      1767
Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys
    190                 195                 200
gac cca gac agt tac tct tgg agg cat ctt cca gtt ggg aaa ctg gag      1815
Asp Pro Asp Ser Tyr Ser Trp Arg His Leu Pro Val Gly Lys Leu Glu
205                 210                 215                 220
tgaactaaac aatatccagg accacaacac catttgggaa ggctgctgca gactggagcc    1875
cacactgaag aatgcaagaa ttattggtga acgaactggc ttccggccag tacgcccca     1935
gattcggcta gaaagagaac agcttcgcac tggaccttca aacacagagg tcatccacaa    1995
ctatggccat ggaggctacg ggctcaccat ccactgggga tgtgccctgg aggcagccaa    2055
gctctttggg agaatcctgg aagaaaagaa attgtccaga atgccaccat cccacctctg    2115
aagactccag tgactgctgc ctccccccac aagaactccc ttctcccctc agccaatgaa    2175
tcaatgtgct ccttcataag ccattgcttc tcctcactt ctttcctcaa agaagcatga     2235
ggtgagagaa agccacaaag tcagtgcctg gagaagggtt cagcccaaca tggggcccct    2295
ctcatcactg aaatccctct accttctctg ggtctggcat tataaagaac agctgaggct    2355
gtcattccat gagtcttcag aagaaaggac agctcagaaa atcaaagagg ccaactgccc    2415
agagccacag aaaatggagg ataattgagg ctaagtaacc tgattacaag ttgtactaac    2475
atattaaagg ttctgaaaag tcctgcagca aagacaacta tctgatgttg tttaacccag    2535
tgcttgctaa acctatctgg ctatggaact cttttgccca gagcacccat gaatgccatg    2595
atacaaatct gagaaaatgc tggaa                                         2620
```

<210> 5
<211> 1576
<212> DNA
<213> Homo sapiens

<220>
<221> 5'UTR
<222> 1 .143

<220>
<221> CDS

<2223> 144..380

<220>
<221> 3'UTR
<222> 381..1576

<400> 5
tgcactccag tccgggctgg cggacagagg gctggaaaca agacgctcca gaatcaggag 60
cttcgcctca ggaaatagca tcctgtgtcc ccgcactgca gttgtctggt ctctccagca 120
gttggtgact tccggctgct gca atg cgt gtg gtg qtg att gga gca gga gtc 173
Met Arg Val Val Val Ile Gly Ala Gly Val
1 5 10
atc ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc 221
Ile Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val
15 20 25
ctg cag cca ctg gac ata aag gtc tac gcg gac cgc ttc acc cca ctc 269
Leu Gln Pro Leu Asp Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu
30 35 40
acc acc acc gac gtg gct gcc ggc ctc tgg cag ccc tac ctt tct gac 317
Thr Thr Thr Asp Val Ala Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp
45 50 55
ccc aac aac cca cag gag gcg acc ctt cct gga agg aca cag ttc tgg 365
Pro Asn Asn Pro Gln Glu Ala Thr Leu Pro Gly Arg Thr Gln Phe Trp
60 65 70
gat ttc gga agc tgaccccag agagctggat atgttcccag attacggcta 417
Asp Phe Gly Ser
75
tggctggttc cacacaagcc taattctgga gggaaagaac tatctacagt ggctgactga 477
aaggttaact gagaggggag tgaagttctt ccagcggaaa gtggagtctt ttgaggaggt 537
ggcaagagaa ggcgcagacg tgattgtcaa ctgcactggg gtatgggctg gggcgctaca 597
acgagacccc ctgctgcagc caggccgggg gcagatcatg aaggtggacg ccccttggat 657
gaagcacttc attctcaccc atgacccaga gagaggcatc tacaattccc cgtacatcat 717
cccagggacc cagacagtta ctcttggagg catcttccag ttgggaaact ggagtgaact 777
aaacaatatc caggaccaca acaccatttg ggaaggctgc tgcagactgg agcccacact 837
gaagaatgca agaattattg gtgaacgaac tggcttccgg ccagtacgcc cccagattcg 897
gctagaaaga gaacagcttc gcactggacc ttcaaaacaca gaggtcatcc acaactatgg 957
ccatggaggc tacgggctca ccatccactg gggatgtgcc ctggaggcag ccaagctctt 1017
tgggagaatc ctggaagaaa agaaattgtc cagaatgcca ccatcccacc tctgaagact 1077
ccagtgactg ctgcctcccc ccacaagaac tcccttctcc cctcagccaa tgaatcaatg 1137
tgctccttca taagccattg cttctccctc acttctttcc tcaaagaagc atgaggtgag 1197
agaaagccac aaagtcagtg cctggagaag ggttcagccc aacatggggc ccctctcatc 1257
actgaaatcc ctctaccttc tctgggtctg gcattataaa gaacagctga ggctgtcatt 1317
ccatgagtct tcagaagaaa ggacagctca gaaaatcaaa gaggccaact gcccagagcc 1377
acagaaaatg gaggataatt gaggctaagt aacctgatta caagttgtac taacatatta 1437
aaggttctga aaagtcctgc agcaaagaca actatctgat gttgtttaac ccagtgcttg 1497
ctaaacctat ctggctatgg aactcttttg cccagagcac ccatgaatgc catgacacaa 1557
atgctgagaa atgctggaa 1576

<210> 6
<211> 1345
<212> DNA
<213> Homo sapiens

<220>
<221> 5'UTR
<222> 1..113

<220>
<221> CDS

<2222> 114..959

<220>
<221> 3'UTR
<222> 960..1345

<220>
<221> polyA_signal
<222> 1301..1326

<400> 6
gaaaccaaag cagcctcctg gattcttccc cgtccctccc tctgtcctgg ggctgtgacc 60
tcctccatgt tattcacagg gtctcagcac gattcatctc aaaggctgct gca atg 116
Met
1
cgt gtg gtg gtg att gga gca gga gtc atc ggg ctg tcc acc gcc ctc 164
Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala Leu
5 10 15
tgc atc cat gag cgc tac cac tca gtc ctg cag cca ctg gac ata aag 212
Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile Lys
20 25 30
gtc tac gcg gac cgc ttc acc cca ctc acc acc acc gac gtg gct gcc 260
Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala Ala
35 40 45
ggc ctc tgg cag ccc tac ctt tct gac ccc aac aac cca cag gag gcg 308
Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu Ala
50 55 60 65
gac tgg agc caa cag acc ttt gac tat ctc ctg agc cat gtc cat tct 356
Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His Ser
70 75 80
ccc aac gct gaa aac ctg ggc ctg ttc cta atc tcg ggc tac aac ctc 404
Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn Leu
85 90 95
ttc cat gaa gcc att ccg gtg gca aga gaa ggc gca gac gtg att gtc 452
Phe His Glu Ala Ile Pro Val Ala Arg Glu Gly Ala Asp Val Ile Val
100 105 110
aac tgc act ggg gta tgg gct ggg gcg cta caa cga gac ccc ctg ctg 500
Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp Pro Leu Leu
115 120 125
cag cca ggc cgg ggg cag atc atg aag gtg gac gcc cct tgg atg aag 548
Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro Trp Met Lys
130 135 140 145
cac ttc att ctc acc cat gac cca gag aga ggc atc tac aat tcc cgg 596
His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr Asn Ser Pro
150 155 160
tac atc atc cca ggg acc cag aca gtt act ctt gga ggc atc ttc cag 644
Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly Ile Phe Gln
165 170 175
ttg gga aac tgg agt gaa cta aac aat atc cag gac cac aac acc att 692
Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His Asn Thr Ile
180 185 190
tgg gaa ggc tgc tgc aga ctg gag ccc aca ctg aag aat gca aga att 740
Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn Ala Arg Ile
195 200 205
att ggt gaa cga act ggc ttc cgg cca gta cgc ccc cag att cgg cta 788
Ile Gly Glu Arg Thr Gly Phe Arg Pro Val Arg Pro Gln Ile Arg Leu
210 215 220 225
gaa aga gaa cag ctt cgc act gga cct tca aac aca gag gtc atc cac 836
Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu Val Ile His

```
                230                 235                 240
aac tat ggc cat gga ggc tac ggg ctc acc atc cac tgg gga tgt gcc      884
Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp Gly Cys Ala
            245                 250                 255
ctg gag gca gcc aag ctc trt ggg aga atc ctg gaa gaa aag aaa ttg      932
Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu Lys Lys Leu
        260                 265                 270
tcc aga atg cca cca tcc cac ctc tgaagactcc agtgactgct gcctcccccc      986
Ser Arg Met Pro Pro Ser His Leu
    275                 280
acaagaactc cctctcccc tcagccaatg aatcaatgtg ctcctcata agccattgct     1046
tctcccttac ttctttcctc aaagaagcat gaggtgagag aaagccacaa agtcagtgcc     1106
tggagaaggg ttcagcccaa catgggggcc ctctcatcac tgaaatcccc ctacctttct     1166
tgggtctggc attataaaga acagctgagg ctgtcattcc atgagtcttc agtagaaagg     1226
acagctcaga aaatcaaaga ggccaactgc ccagagccac agaaaatgga ggataattga     1286
ggctaagtaa cctgattaca agttgtacta acatattaaa ggttctgaaa agtcctgca      1345

<210> 7
<211> 347
<212> PRT
<213> Homo sapiens

<400> 7
Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
1               5                  10                  15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
            20                  25                  30
Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
        35                  40                  45
Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
    50                  55                  60
Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His
65                  70                  75                  80
Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn
                85                  90                  95
Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys Asp Thr Val Leu
            100                 105                 110
Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met Phe Pro Asp Tyr
        115                 120                 125
Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu Gly Lys Asn Tyr
    130                 135                 140
Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly Val Lys Phe Phe
145                 150                 155                 160
Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg Glu Gly Ala Asp
                165                 170                 175
Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp
            180                 185                 190
Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro
        195                 200                 205
Trp Met Lys His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr
    210                 215                 220
Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly
225                 230                 235                 240
Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His
                245                 250                 255
Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn
            260                 265                 270
Ala Arg Ile Ile Gly Glu Arg Thr Gly Phe Arg Pro Val Arg Pro Gln
        275                 280                 285
```

Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu
290 295 300
Val Ile His Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp
305 310 315 320
Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu
325 330 335
Lys Lys Leu Ser Arg Met Pro Pro Ser His Leu
340 345

<210> 8
<211> 220
<212> PRT
<213> Homo sapiens

<400> 8
Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
1 5 10 15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
20 25 30
Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
35 40 45
Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
50 55 60
Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His
65 70 75 80
Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn
85 90 95
Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys Asp Thr Val Leu
100 105 110
Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met Phe Pro Asp Tyr
115 120 125
Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu Gly Lys Asn Tyr
130 135 140
Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly Val Lys Phe Phe
145 150 155 160
Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg Glu Gly Ala Asp
165 170 175
Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp
180 185 190
Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys Asp Pro Asp Ser
195 200 205
Tyr Ser Trp Arg His Leu Pro Val Gly Lys Leu Glu
210 215 220

<210> 9
<211> 78
<212> PRT
<213> Homo sapiens

<400> 9
Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
1 5 10 15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
20 25 30
Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
35 40 45
Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
50 55 60
Ala Thr Leu Pro Gly Arg Thr Gln Phe Trp Asp Phe Gly Ser

65 70 75

<210> 10
<211> 281
<212> PRT
<213> Homo sapiens

<400> 10
Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
1 5 10 15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
20 25 30
Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
35 40 45
Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
50 55 60
Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His
65 70 75 80
Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn
85 90 95
Leu Phe His Glu Ala Ile Pro Val Ala Arg Glu Gly Ala Asp Val Ile
100 105 110
Val Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp Pro Leu
115 120 125
Leu Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro Trp Met
130 135 140
Lys His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr Asn Ser
145 150 155 160
Pro Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly Ile Phe
165 170 175
Gln Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His Asn Thr
180 185 190
Ile Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn Ala Arg
195 200 205
Ile Ile Gly Glu Arg Thr Gly Phe Arg Pro Val Arg Pro Gln Ile Arg
210 215 220
Leu Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu Val Ile
225 230 235 240
His Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp Gly Cys
245 250 255
Ala Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu Lys Lys
260 265 270
Leu Ser Arg Met Pro Pro Ser His Leu
275 280

<210> 11
<211> 456
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> 152
<223> 99-16105-152 : polymorphic base A or G

<220>
<221> misc_binding
<222> 140..164
<223> 99-16105-152.probe

<220>
<221> primer_bind
<222> 133..151
<223> 99-16105-152.mis

<220>
<221> primer_bind
<222> 153..171
<223> 99-16105-152.mis complement

<220>
<221> primer_bind
<222> 1..19
<223> 99-16105.pu

<220>
<221> primer_bind
<222> 437..456
<223> 99-16105.rp complement

<400> 11
cgctttgttg tattctttgt tatttatcca ttttgccaaa ttatctgcaa gtagaaatat 60
cgaaataaga agctctttag caatttactt tggatattgg ttttcttttg aaggacagtt 120
attaaaatag cttgtaggat tactcatttt crttttttctt ctttttaaat ataaagcaat 180
gtcatcactt ttttccctgt attatatttc tcctcaataa ttgatatgct acattaaagg 240
aacacaaaat ggtcttaatt atgcaataat gatcaaggca aagagtgttt cctgggaact 300
aatggttgcc tgagaggagg tgatggcttg aggtccagct ggttattaag ccgcaggaaa 360
tgctgcaggc caagatttgt attatttctc tgagatgaaa atgaacccaa aaaaaggcaa 420
aatgggtttt tctccactaa tgggtaaaat gaactc 456

<210> 12
<211> 463
<212> DNA
<213> Homo sapiens

<220>
<221> allele
<222> 215
<223> 99-5919-215 : polymorphic base A or G

<220>
<221> misc_binding
<222> 203..227
<223> 99-5919-215.probe

<220>
<221> primer_bind
<222> 196..214
<223> 99-5919-215.mis

<220>
<221> primer_bind
<222> 216..234
<223> 99-5919-215.mis complement

<220>
<221> primer_bind
<222> 11..29

<223> 99-5919.pu

<220>
<221> primer_bind
<222> 445..465
<223> 99-5919.rp complement

<400> 13
[illegible]ctgga ctacagcaat gcagatttca attctgccat tgaattccca gacatattcg 60
[illegible]tcat ttcatcccc caccaccctg ccatttcctt cgtgttaact tgttttcctg 120
agtccacaaa atcacctttt cctgtataca ttttaggat gtcagacttt attctaatga 180
ttttcccttag ttgcccccca aaattgtatt ctacrgtgtg attttaaagc tgaattttca 240
aaatgatatt tcatatctat attttcacaa gcttttcttc tatgaatgtt attgtcagct 300
gtcaagggtgt gagatggtac ttgatactac attctttcca agctgttgcc tgaatcggtt 360
taaaacaaag tcattactag gctgtaaact gttgctctgc aaaattgagc agcacgtatt 420
taaccactca tacttcttag ctctccaaca ctttgagtcr ata 463

<210> 13
<211> 742
<212> DNA
<213> Homo sapiens

<220>
<221> 5'UTR
<222> 1..46

<220>
<221> CDS
<222> 47..508

<220>
<221> 3'UTR
<222> 509..742

<220>
<221> polyA_signal
<222> 718..723

<220>
<221> allele
<222> 21
<223> 8-135 112 : polymorphic base C or T

<220>
<221> allele
<222> 75
<223> 8-135-166 : polymorphic base A or C

<220>
<221> allele
<222> 135
<223> 99-16038-118 : polymorphic base A or G

<220>
<221> allele
<222> 194
<223> 8-137-152 : polymorphic base G or T

<220>

<220>
<221> allele
<222> 224
<223> 8-137-182 : polymorphic base A or G

<220>
<221> allele
<222> 390
<223> 8-130-220 : polymorphic base A or C

<220>
<221> allele
<222> 406
<223> 8-130-236 : polymorphic base A or G

<220>
<221> allele
<222> 578
<223> 8-131-199 : polymorphic base A or C

<220>
<221> allele
<222> 641
<223> 8-132-97 : polymorphic base C or T

<220>
<221> allele
<222> 708
<223> 8-132-164 : polymorphic base C or T

<220>
<221> allele
<222> 723
<223> 8-132-179 : polymorphic base A or T

<220>
<221> VARIANT
<222> 75
<223> Xaa=Ser or Tyr

<220>
<221> VARIANT
<222> 135
<223> Xaa=Lys or Arg

<220>
<221> VARIANT
<222> 194
<223> Xaa=Glu or Stop

<220>
<221> VARIANT
<222> 224
<223> Xaa=Gly or Arg

<220>
<221> VARIANT
<222> 490
<223> Xaa=Ala or Asp

<400> 13
tcatctctgc ttcacaatgc ygatgattta gctgggagga cccaaa atg ctg gaa 55
Met Leu Glu
1
aag ctg atg ggt gct gat tmt ctc cag ctt ttc aga tcc aga tat aca 103
Lys Leu Met Gly Ala Asp Xaa Leu Gln Leu Phe Arg Ser Arg Tyr Thr
5 10 15
ttg ggt aaa atc tac ttc ata ggt ttt caa arg agc att ctt ctg agc 151
Leu Gly Lys Ile Tyr Phe Ile Gly Phe Gln Xaa Ser Ile Leu Leu Ser
20 25 30 35
aaa tct gaa aac tct cta aac tct att gca aag gag aca gaa kaa gga 199
Lys Ser Glu Asn Ser Leu Asn Ser Ile Ala Lys Glu Thr Glu Xaa Gly
40 45 50
aga gag acg gta aca agg aaa gaa rga tgg aag aga agg cat gag gac 247
Arg Glu Thr Val Thr Arg Lys Glu Xaa Trp Lys Arg Arg His Glu Asp
55 60 65
ggc tat ttg gaa atg gca cag agg cat tta cag aga tca tta tgt cct 295
Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser Leu Cys Pro
70 75 80
tgg gtc tct tac ctt cct cag ccc tat gca gag ctt gaa gaa gta agc 343
Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu Glu Val Ser
85 90 95
agc cat gtt gga aaa gtc ttc atg gca aga aac tat gag ttc ctt gmc 391
Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu Phe Leu Xaa
100 105 110 115
tat gag gcc tct aar gac cgc agg cag cct cta gaa cga atg tgg acc 439
Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg Met Trp Thr
120 125 130
tgc aac tac aac cag caa aaa gac cag tca tgc aac cac aag gaa ata 487
Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His Lys Glu Ile
135 140 145
act tct acc aaa gct gaa tgagtttgga agcagattct tcccagccaa 535
Thr Ser Thr Lys Ala Glu
150
tccttctgat gacaatgtag tctggccaac atcttcactg gamtctgacg gactctgtgt 595
ctgggaccca gctgataaca cgtggtgatg ggattgtatt tgcaaytctc tggtcagtaa 655
gtgataaaat gccatttcta tgcacccacc tggcctgtgt gactgggaga atytctcttt 715
ttattaawtg tgcttcaagt tttaaca 742

<210> 14
<211> 153
<212> PRT
<213> Homo sapiens

<220>
<221> VARIANT
<222> 10
<223> Xaa=Ser or Tyr

<220>
<221> VARIANT
<222> 30
<223> Xaa=Lys or Arg

<220>
<221> VARIANT
<222> 50
<223> Xaa=Glu or Stop

<220>
<221> VARIANT
<222> 60
<223> Xaa=Gly or Arg

<220>
<221> VARIANT
<222> 115
<223> Xaa=Ala or Asp

<220>
<221> SITE
<222> 63..64
<223> basic protease cleavage site

<220>
<221> SITE
<222> 122..123
<223> basic protease cleavage site

<220>
<221> DISULFID
<222> 132..142

<220>
<221> PEPTIDE
<222> 1..64

<220>
<221> PEPTIDE
<222> 65..123

<220>
<221> PEPTIDE
<222> 124..153

<400> 14
Met Leu Glu Lys Leu Met Gly Ala Asp Xaa Leu Gln Leu Phe Arg Ser
1 5 10 15
Arg Tyr Thr Leu Gly Lys Ile Tyr Phe Ile Gly Phe Gln Xaa Ser Ile
20 25 30
Leu Leu Ser Lys Ser Glu Asn Ser Leu Asn Ser Ile Ala Lys Glu Thr
35 40 45
Glu Xaa Gly Arg Glu Thr Val Thr Arg Lys Glu Xaa Trp Lys Arg Arg
50 55 60
His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser
65 70 75 80
Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu
85 90 95
Gln Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu
100 105 110
Phe Leu Xaa Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg
115 120 125
Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His
130 135 140
Lys Glu Ile Thr Ser Thr Lys Ala Glu
145 150

<210> 15

<211> 476
<212> DNA
<213> Homo sapiens

<400> 15
cat gag gac ggc tat ttg gaa atg gca cag agg cat tta cag aga tca 48
His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser
1 5 10 15
tta tgt cct tgg gtc tct tac ctt cct cag ccc tat gca gag ctt gaa 96
Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu
20 25 30
gaa gta agc agc cat gtt gga aaa gtc ttc atg gca aga aac tat gag 144
Glu Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu
35 40 45
ttc ctt gcc tat gag gcc tct aag gac cgc agg cag cct cta gaa cga 192
Phe Leu Ala Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg
50 55 60
arg tgg acc tgc aac tac aac cag caa aaa gac cag tca tgc aac cac 240
Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His
65 70 75 80
aag gaa ata act tct acc aaa gct gaa tgagtttgga agcagattct 287
Lys Glu Ile Thr Ser Thr Lys Ala Glu
85
tccagccaa tcttctgat gacaatgtag tctggccaac atcttcactg gactctgacg 347
gactctgtgt ctgggaccca gctgataaca cgtggtgatg ggattgtatt tgcaactctc 407
tggtcagtaa gtgataaaat gccatttcta tgcacccacc tggcctgtgt gactgggaga 467
atctctctt 476

<210> 16
<211> 89
<212> PRT
<213> Homo sapiens

<400> 16
His Glu Asp Gly Tyr Leu Glu Met Ala Gln Arg His Leu Gln Arg Ser
1 5 10 15
Leu Cys Pro Trp Val Ser Tyr Leu Pro Gln Pro Tyr Ala Glu Leu Glu
20 25 30
Glu Val Ser Ser His Val Gly Lys Val Phe Met Ala Arg Asn Tyr Glu
35 40 45
Phe Leu Ala Tyr Glu Ala Ser Lys Asp Arg Arg Gln Pro Leu Glu Arg
50 55 60
Met Trp Thr Cys Asn Tyr Asn Gln Gln Lys Asp Gln Ser Cys Asn His
65 70 75 80
Lys Glu Ile Thr Ser Thr Lys Ala Glu
85

<210> 17
<211> 1533
<212> DNA
<213> Homo sapiens

<400> 17
ttgggctcca ttgcaacccg aggcgagact agagttccca agcgagaagg gaagagggag 60
tgggtgcacg tggaaggcgg acagagggct ggaaacaaga cgctccagaa tcaggagctt 120
cccctcagga aatagcatcc tgtgtccccg cactgcagtt gtctggtctc tccagcagtt 180
tggtacttcc ggctgctgca atg cgt gtg gtg gtg att gga gca gga gtc atc 233
Met Arg Val Val Val Ile Gly Ala Gly Val Ile
1 5 10

```
ggg ctg tcc acc gcc ctc tgc atc cat gag cgc tac cac tca gtc ctg      281
Gly Leu Ser Thr Ala Leu Cys Ile His Glu Arg Tyr His Ser Val Leu
            15                  20                  25
cag cca ctg cac ata aag gtc tac gcg gac cgc ttc acc gca ctc acc      329
Gln Pro Leu His Ile Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr
        30                  35                  40
acc acc gac gtg gct gcc gac ctc tgg cag ccc tac ctt tct gac ccc      377
Thr Thr Asp Val Ala Ala Asp Leu Trp Gln Pro Tyr Leu Ser Asp Pro
    45                  50                  55
aac aac cca cag gag gtg gac tgg agc caa cag acc ttt gac tat ctc      425
Asn Asn Pro Gln Glu Val Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu
60                  65                  70                  75
ctg agc cat gtc cat tct ccc aac gct gaa aac ctg ggc ctg ttc cta      473
Leu Ser His Val His Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu
                80                  85                  90
atc tcg ggc tac aac ctc ttc cat gaa gcc att ccg gac cct tcc tgg      521
Ile Ser Gly Tyr Asn Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp
            95                  100                 105
aag gac aca gtt ctg gga ttt cgg aag ctg acc ccc aga gag ctg gat      569
Lys Asp Thr Val Leu Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp
        110                 115                 120
atg ttc cca gat tac ggc tat ggc tgg ttc cac aca agc cta att ctg      617
Met Phe Pro Asp Tyr Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu
    125                 130                 135
gag gga aag aac tat cta cag tgg ctg act gaa agg tta act gag agg      665
Glu Gly Lys Asn Tyr Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg
140                 145                 150                 155
gga gtg aag ttc ttc cag cgg aaa gtg gag tct ttt gag gag gtg gca      713
Gly Val Lys Phe Phe Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala
                160                 165                 170
aga gaa ggc gca gac gtg att gtc aac tgc act ggg gta tgg gct ggg      761
Arg Glu Gly Ala Asp Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly
            175                 180                 185
gcg cta caa cga gac ccc ctg ctg cag cca ggc cgg ggg cag atc atg      809
Ala Leu Gln Arg Asp Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met
        190                 195                 200
aag gtg gac gcc cct tgg atg aag cac ttc att ctc acc cat gac cca      857
Lys Val Asp Ala Pro Trp Met Lys His Phe Ile Leu Thr His Asp Pro
    205                 210                 215
gag aga ggc atc tac aat tcc ccg tac atc atc cca ggg acc cag aca      905
Glu Arg Gly Ile Tyr Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln Thr
220                 225                 230                 235
gtt act ctt gga ggc atc ttc cag ttg gga aac tgg agt gaa cta aac      953
Val Thr Leu Gly Gly Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu Asn
                240                 245                 250
aat atc cag gac cac aac acc att tgg gaa ggc tgc tgc aga ctg gag     1001
Asn Ile Gln Asp His Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu Glu
            255                 260                 265
ccc aca ctg aag aat gca aga att att ggt gaa gca act ggc ttc cgg     1049
Pro Thr Leu Lys Asn Ala Arg Ile Ile Gly Glu Ala Thr Gly Phe Arg
        270                 275                 280
cca gta cgc ccc cag att cgg cta gaa aga gaa cag ctt cgc act gga     1097
Pro Val Arg Pro Gln Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr Gly
    285                 290                 295
cct tca aac aca gag gtc atc cac aac tat ggc cat gga ggc tac ggg     1145
Pro Ser Asn Thr Glu Val Ile His Asn Tyr Gly His Gly Gly Tyr Gly
300                 305                 310                 315
ctc acc atc cac tgg gga tgt gcc ctg gag gca gcc aag ctc ttt ggg     1193
Leu Thr Ile His Trp Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe Gly
```

```
                320                 325                 330
aga atc ctg gaa gaa aag aaa ttg tcc aga atg cca cca tcc cac ctc     1241
Arg Ile Leu Glu Glu Lys Lys Leu Ser Arg Met Pro Pro Ser His Leu
            335                 340                 345
tgaagactcc agtgactgct gcctcccccc acaagaactc cctctcccca tcagccaatg   1301
aatcaatgtg ctccttcata agccattgct tctcctcac ttttttcctc aaagaagcat    1361
gaggtgagag aaagccacra agtcagtgcc tggagaaagg ttcagcccaa catggggccc   1421
ctctcatcac tgaaatccct ctacctttcc tgggtctggc attataaaga acagctgagg   1481
gtgtcattcc atgagtcttc agaagaaagg acagctcaga aagtcaaaga ggccaactgc   1541
ccaaagccac aqaaaaatgga qqataattga qqctaagtaa cctgattaca agttgtacta   1601
acatattaaa ggttctgaaa agtcctgcaa aa                                 1633

<210> 18
<211> 347
<212> PRT
<213> Homo sapiens

<400> 18
Met Arg Val Val Val Ile Gly Ala Gly Val Ile Gly Leu Ser Thr Ala
1               5                   10                  15
Leu Cys Ile His Glu Arg Tyr His Ser Val Leu Gln Pro Leu Asp Ile
            20                  25                  30
Lys Val Tyr Ala Asp Arg Phe Thr Pro Leu Thr Thr Thr Asp Val Ala
        35                  40                  45
Ala Gly Leu Trp Gln Pro Tyr Leu Ser Asp Pro Asn Asn Pro Gln Glu
    50                  55                  60
Ala Asp Trp Ser Gln Gln Thr Phe Asp Tyr Leu Leu Ser His Val His
65                  70                  75                  80
Ser Pro Asn Ala Glu Asn Leu Gly Leu Phe Leu Ile Ser Gly Tyr Asn
                85                  90                  95
Leu Phe His Glu Ala Ile Pro Asp Pro Ser Trp Lys Asp Thr Val Leu
            100                 105                 110
Gly Phe Arg Lys Leu Thr Pro Arg Glu Leu Asp Met Phe Pro Asp Tyr
        115                 120                 125
Gly Tyr Gly Trp Phe His Thr Ser Leu Ile Leu Glu Gly Lys Asn Tyr
    130                 135                 140
Leu Gln Trp Leu Thr Glu Arg Leu Thr Glu Arg Gly Val Lys Phe Phe
145                 150                 155                 160
Gln Arg Lys Val Glu Ser Phe Glu Glu Val Ala Arg Glu Gly Ala Asp
                165                 170                 175
Val Ile Val Asn Cys Thr Gly Val Trp Ala Gly Ala Leu Gln Arg Asp
            180                 185                 190
Pro Leu Leu Gln Pro Gly Arg Gly Gln Ile Met Lys Val Asp Ala Pro
        195                 200                 205
Trp Met Lys His Phe Ile Leu Thr His Asp Pro Glu Arg Gly Ile Tyr
    210                 215                 220
Asn Ser Pro Tyr Ile Ile Pro Gly Thr Gln Thr Val Thr Leu Gly Gly
225                 230                 235                 240
Ile Phe Gln Leu Gly Asn Trp Ser Glu Leu Asn Asn Ile Gln Asp His
                245                 250                 255
Asn Thr Ile Trp Glu Gly Cys Cys Arg Leu Glu Pro Thr Leu Lys Asn
            260                 265                 270
Ala Arg Ile Ile Gly Glu Ala Thr Gly Phe Arg Pro Val Arg Pro Gln
        275                 280                 285
Ile Arg Leu Glu Arg Glu Gln Leu Arg Thr Gly Pro Ser Asn Thr Glu
    290                 295                 300
Val Ile His Asn Tyr Gly His Gly Gly Tyr Gly Leu Thr Ile His Trp
305                 310                 315                 320
Gly Cys Ala Leu Glu Ala Ala Lys Leu Phe Gly Arg Ile Leu Glu Glu
```

```
                325                 330                 335
Lys Lys Leu Ser Arg Met Pro Pro Ser His Leu
            340                 345

<210> 19
<211> 1200
<212> DNA
<213> Homo sapiens

<400> 19
atg gac aca gca cgg att gca gtt gtc ggg gca ggt gtg gtg ggg ctc      48
Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu
1               5                   10                  15
tcc acg gct gtg tgc atc tcc aaa ctg gtg ccc cga tgc tcc gtt acc      96
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr
            20                  25                  30
atc att tca gac aag ttt act cca gat tcc acc agt gat gtg gca gcc     144
Ile Ile Ser Asp Lys Phe Thr Pro Asp Ser Thr Ser Asp Val Ala Ala
        35                  40                  45
gga atg ctt att cct cac act tat cca gat aca ccc att cac acg cag     192
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln
    50                  55                  60
aag cag tgg ttc aga gaa acc ttt aat cac ctc ttt gca att gcc aat     240
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn
65                  70                  75                  80
tct gca gaa gct gga gat gct ggt gtt cat ttg gta tca ggt tgg cag     288
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Trp Gln
                85                  90                  95
ata ttt cag agc act ccg act gaa gaa gtg cca ttc tgg gct gac gtg     336
Ile Phe Gln Ser Thr Pro Thr Glu Glu Val Pro Phe Trp Ala Asp Val
            100                 105                 110
gtt ctg gga ttt cga aag atg act gag gct gag ctg aag gaa ttc ccc     384
Val Leu Gly Phe Arg Lys Met Thr Glu Ala Glu Leu Lys Lys Phe Pro
        115                 120                 125
cag tat gtg ttt ggt cag gct ttt aca acc ctg aaa tgt gaa tgc cct     432
Gln Tyr Val Phe Gly Gln Ala Phe Thr Thr Leu Lys Cys Glu Cys Pro
    130                 135                 140
gcc tac ctc ccg tgg ttg gag aaa agg ata aag gga agt gga ggc tgg     480
Ala Tyr Leu Pro Trp Leu Glu Lys Arg Ile Lys Gly Ser Gly Gly Trp
145                 150                 155                 160
aca ctc act cgg cga ata gaa gac ctg tgg gaa ctt cat ccg tcc ttt     528
Thr Leu Thr Arg Arg Ile Glu Asp Leu Trp Glu Leu His Pro Ser Phe
                165                 170                 175
gac atc gtg gtc aac tgt tca ggc ctt gga agc aga cag ctt gca gga     576
Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser Arg Gln Leu Ala Gly
            180                 185                 190
gac tca aag att ttc cct gta agg ggc caa gtc ctc caa gtt cag gct     624
Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val Leu Gln Val Gln Ala
        195                 200                 205
ccc tgg gtg gag cat ttt atc cga gat ggc agt ggg ctg aca tat att     672
Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser Gly Leu Thr Tyr Ile
    210                 215                 220
tat cct ggt aca tcc cat gta acc cta ggt gga act agg caa aaa ggg     720
Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly Thr Arg Gln Lys Gly
225                 230                 235                 240
gac tgg aat ctg tcc ccg gat gca gaa aat agc aga gag att ctt tcc     768
Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser Arg Glu Ile Leu Ser
                245                 250                 255
cga tgc tgt gct ctg gag ccc tcc ctc caa gga gcc tgc aac atc agg     816
```

```
Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly Ala Cys Asn Ile Arg
            260                 265                 270
gag aag gtg ggc ttg agg ccc tac agg cca ggc gtg cga ctg cag aca      864
Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly Val Arg Leu Gln Thr
        275                 280                 285
gag ctc ttt ggg cga gat gga cag agg ctg cct gta gtc cac cac tat      912
Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro Val Val His His Tyr
    290                 295                 300
ggc cat ggg agt ggg ggc atc tca gtg cac tgg ggc act gct ctg gag      960
Gly His Gly Ser Gly Gly Ile Ser Val His Trp Gly Thr Ala Leu Glu
305                 310                 315                 320
gcc gcc agg ctg gtg agc gag tgt gtc cat gcc ctc agg acc ccc att     1008
Ala Ala Arg Leu Val Ser Glu Cys Val His Ala Leu Arg Thr Pro Ile
                325                 330                 335
ccc aag tca aac ctg tagatgacat aaaatgacag caaagagact gagagactgt     1063
Pro Lys Ser Asn Leu
            340
tgatcaaagc acagaacagg ttcaaataac ttttccactg catgaaagtt taattagaca    1123
tttctttgtt ttcaacatta gaagtggtgt aacatgtaag ctgagcacgg tagcatgcct    1183
atagtcccag ctacttg                                                  1200

<210> 20
<211> 1023
<212> DNA
<213> Homo sapiens

<400> 20
atg gac aca gca cgg att gca gtt gtc ggg gca ggt gtg gtg ggg ctc       48
Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu
1               5                   10                  15
tcc acg gct gtg tgc atc tcc aaa ctg gtg ccc cga tgc tcc gtt acc       96
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr
            20                  25                  30
atc att tca gac aag ttt act cca gat acc acc agt gat gtg gca gcc      144
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala
        35                  40                  45
gga atg ctt att cct cac act tat cca gat aca ccc att cac acg cag      192
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln
    50                  55                  60
aag cag tgg ttc aga gaa acc ttt aat cac ctc ttt gca att gcc aat      240
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn
65                  70                  75                  80
tct gca gaa gct gga gat gct ggt gtt cat ttg gta tca ggg ata aag      288
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Ile Lys
                85                  90                  95
gga agt gga ggc tgg aca ctc act cgg cga ata gaa gac ctg tgg gaa      336
Gly Ser Gly Gly Trp Thr Leu Thr Arg Arg Ile Glu Asp Leu Trp Glu
            100                 105                 110
ctt cat ccg tcc ttt gac atc gtg gtc aac tgt tca ggc ctt gga agc      384
Leu His Pro Ser Phe Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser
        115                 120                 125
aga cag ctt gca gga gac tca aag att ttc cct gta agg ggc caa gtc      432
Arg Gln Leu Ala Gly Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val
    130                 135                 140
ctc caa gtt cag gct ccc tgg gtg gag cat ttt atc cga gat ggc agt      480
Leu Gln Val Gln Ala Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser
145                 150                 155                 160
ggg ctg aca tat att tat cct ggt aca tcc cat gta acc cta ggt gga      528
Gly Leu Thr Tyr Ile Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly
```

```
                    165                 170                 175
act agg caa aaa ggg gac tgg aat ctg tcc ccg gat gca gaa aat agc      576
Thr Arg Gln Lys Gly Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser
            180                 185                 190
aga gag att ctt tct cga tgc tgt gct ctg gag ccc tcc ctc cac gga      624
Arg Glu Ile Leu Ser Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly
        195                 200                 205
gcc tgc aac atc agg gag aag gtg ggc ttg agg ccc tac agg cca ggc      672
Ala Cys Asn Ile Arg Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly
    210                 215                 220
gtg cga ctg cag aca gag ctc ctt gcg cga gat gga cag agg ctg cct      720
Val Arg Leu Gln Thr Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro
225                 230                 235                 240
gta gtc cac cac tat ggc cat ggg agt ggg ggc atc tca gtg cac tgg      768
Val Val His His Tyr Gly His Gly Ser Gly Gly Ile Ser Val His Trp
                245                 250                 255
ggc act gct ctg gag gcc gcc agg ctg gtg agc gag tgt gtc cat gcc      816
Gly Thr Ala Leu Glu Ala Ala Arg Leu Val Ser Glu Cys Val His Ala
            260                 265                 270
ctc agg acc ccc att ccc aag tca aac ctg tagatgacat aaaatgacag       866
Leu Arg Thr Pro Ile Pro Lys Ser Asn Leu
        275                 280
caaagagact gagagactgt tgatcaaagc acagaacagg ttcaaataac ttttcactg    926
catgaaagtt taattagaca tttctttgtt ttcaacatta gaagtggtgt aacatgtaag    986
ctgagcacgg tagcatgcct atagtcccag ctatttg                           1023

<210> 21
<211> 341
<212> PRT
<213> Homo sapiens

<400> 21
Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu
1               5                   10                  15
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr
            20                  25                  30
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala
        35                  40                  45
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln
    50                  55                  60
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn
65                  70                  75                  80
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Trp Gln
                85                  90                  95
Ile Phe Gln Ser Thr Pro Thr Glu Glu Val Pro Phe Trp Ala Asp Val
            100                 105                 110
Val Leu Gly Phe Arg Lys Met Thr Glu Ala Glu Leu Lys Lys Phe Pro
        115                 120                 125
Gln Tyr Val Phe Gly Gln Ala Phe Thr Thr Leu Lys Cys Glu Cys Pro
    130                 135                 140
Ala Tyr Leu Pro Trp Leu Glu Lys Arg Ile Lys Gly Ser Gly Gly Trp
145                 150                 155                 160
Thr Leu Thr Arg Arg Ile Glu Asp Leu Trp Glu Leu His Pro Ser Phe
                165                 170                 175
Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser Arg Gln Leu Ala Gly
            180                 185                 190
Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val Leu Gln Val Gln Ala
        195                 200                 205
Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser Gly Leu Thr Tyr Ile
```

```
    210                 215                 220
Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly Thr Arg Gln Lys Gly
225                 230                 235                 240
Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser Arg Glu Ile Leu Ser
                245                 250                 255
Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly Ala Cys Asn Ile Arg
            260                 265                 270
Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly Val Arg Leu Gln Thr
        275                 280                 285
Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro Val Val His His Tyr
    290                 295                 300
Gly His Gly Ser Gly Gly Ile Ser Val His Trp Gly Thr Ala Leu Glu
305                 310                 315                 320
Ala Ala Arg Leu Val Ser Glu Cys Val His Ala Leu Arg Thr Pro Ile
                325                 330                 335
Pro Lys Ser Asn Leu
            340
```

<210> 22
<211> 282
<212> PRT
<213> Homo sapiens

```
<400> 22
Met Asp Thr Ala Arg Ile Ala Val Val Gly Ala Gly Val Val Gly Leu
1               5                   10                  15
Ser Thr Ala Val Cys Ile Ser Lys Leu Val Pro Arg Cys Ser Val Thr
            20                  25                  30
Ile Ile Ser Asp Lys Phe Thr Pro Asp Thr Thr Ser Asp Val Ala Ala
        35                  40                  45
Gly Met Leu Ile Pro His Thr Tyr Pro Asp Thr Pro Ile His Thr Gln
    50                  55                  60
Lys Gln Trp Phe Arg Glu Thr Phe Asn His Leu Phe Ala Ile Ala Asn
65                  70                  75                  80
Ser Ala Glu Ala Gly Asp Ala Gly Val His Leu Val Ser Gly Ile Lys
                85                  90                  95
Gly Ser Gly Gly Trp Thr Leu Thr Arg Arg Ile Glu Asp Leu Trp Glu
            100                 105                 110
Leu His Pro Ser Phe Asp Ile Val Val Asn Cys Ser Gly Leu Gly Ser
        115                 120                 125
Arg Gln Leu Ala Gly Asp Ser Lys Ile Phe Pro Val Arg Gly Gln Val
    130                 135                 140
Leu Gln Val Gln Ala Pro Trp Val Glu His Phe Ile Arg Asp Gly Ser
145                 150                 155                 160
Gly Leu Thr Tyr Ile Tyr Pro Gly Thr Ser His Val Thr Leu Gly Gly
                165                 170                 175
Thr Arg Gln Lys Gly Asp Trp Asn Leu Ser Pro Asp Ala Glu Asn Ser
            180                 185                 190
Arg Glu Ile Leu Ser Arg Cys Cys Ala Leu Glu Pro Ser Leu His Gly
        195                 200                 205
Ala Cys Asn Ile Arg Glu Lys Val Gly Leu Arg Pro Tyr Arg Pro Gly
    210                 215                 220
Val Arg Leu Gln Thr Glu Leu Leu Ala Arg Asp Gly Gln Arg Leu Pro
225                 230                 235                 240
Val Val His His Tyr Gly His Gly Ser Gly Gly Ile Ser Val His Trp
                245                 250                 255
Gly Thr Ala Leu Glu Ala Ala Arg Leu Val Ser Glu Cys Val His Ala
            260                 265                 270
Leu Arg Thr Pro Ile Pro Lys Ser Asn Leu
```

275 280

<210> 23
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide 24-1443-126

<220>
<221> allele
<222> 24
<223> polymorphic base C or T

<400> 23
taaggcttag taagttggag nacyaggatc agaagacagg tctgcct 47

<210> 24
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide 24-1457-52

<220>
<221> allele
<222> 24
<223> polymorphic base C or A

<400> 24
tctgagatgc ccctgtgtcc tctmagggag tagtggctga gcatttc 47

<210> 25
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide 27-93-181

<220>
<221> allele
<222> 24
<223> polymorphic base C or T

<400> 25
ccccagcctctg ccactggcga rctytgtggc cttgggcaag ttactcc 47

<210> 26
<211> 47
<212> DNA
<213> Artificial Sequence

<220>
<223> oligonucleotide 24-1461-256

<220>

<221> allele
<222> 24
<223> polymorphic base A or G

<400> 26
gatggctctg gcattttcag ggarcagtca tgtctgatct caagttc 47
1